TU Delft

Delft University of Technology

Document Version
Final published version

Citation (APA)
Dai, Z., Li, D., Feng, Y., & Yang, C. (2026). Decision evolution and heterogeneity aware pedestrian wayfinding behaviour modelling in VR integrated transportation hub. *Transportation Research Part C: Emerging Technologies*, *186*, Article 105595. https://doi.org/10.1016/j.trc.2026.105595

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Transportation Research Part C 186 (2026) 105595

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Decision evolution and heterogeneity aware pedestrian wayfinding behaviour modelling in VR integrated transportation hub

Zhicheng Dai [a,b], Dewei Li [a,c,*], Yan Feng [d], Chenyi Yang [a]

[a] School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China
[b] Department of Built Environment, Eindhoven University of Technology, 5612 AP, Eindhoven, The Netherlands
[c] Frontiers Science Center for Smart High-speed Railway System, Beijing Jiaotong University, Beijing, 100044, China
[d] Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2600 AA, Delft, The Netherlands

ARTICLE INFO

Keywords:
Virtual reality
Decision cascade
Wayfinding
Visual information
Pedestrian heterogeneity

ABSTRACT

The dynamic evolution and heterogeneity of passenger wayfinding decisions in integrated transport hubs have a significant impact on both operational efficiency and user experience. However, existing static models fall short in capturing the temporal variability of passengers' cognitive states in response to environmental and situational changes. This study develops a virtual reality scenario of an integrated transport hub and conducts non-immersive behavioral experiments to support the construction of a novel dynamic modeling framework–**D**ynamic **H**idden **M**arkov **M**odel-Logit (DHMM-Logit), which integrates a multi-state hidden Markov model with a Logit model. For the first time, decision cascading analysis is introduced into this framework, utilizing mutual information theory to uncover the temporal dependency and decay mechanism of historical decisions on current choices. These insights guide both the hyperparameter setting and discretization of decision sequences in the DHMM-Logit model. The framework comprehensively incorporates spatial syntax metrics, the use of 2D navigation tools and travel purposes to account for spatial and individual heterogeneity. In addition, a graph embedding-based high-order semantic encoding of nodes is introduced as explanatory variables, enhancing the model's ability to fit and generalize sequential pedestrian decision-making processes. Empirical validation in the Shanghai Hongqiao Integrated Transport Hub demonstrates that the proposed DHMM-Logit model significantly outperforms baseline methods. The findings reveal pronounced latent cognitive state transitions during pedestrian wayfinding, with travel mode and navigation usage exerting significant influence on passengers' spatial sensitivity and cognitive processes. This research provides a solid theoretical and empirical foundation for the optimization of hub spatial design and the implementation of personalized information guidance strategies.

1. Introduction

1.1. Research background and problem statement

In the era of rapid urbanization and increasingly sophisticated multimodal transportation systems, integrated transport hubs, such as high-speed railway stations and airports, are becoming critical nodes within urban transport networks, playing a key role

* Corresponding author.
E-mail addresses: 21114047@bjtu.edu.cn (Z. Dai), lidw@bjtu.edu.cn (D. Li), y.feng@tudelft.nl (Y. Feng), 24120855@bjtu.edu.cn (C. Yang).

https://doi.org/10.1016/j.trc.2026.105595
Received 8 August 2025; Received in revised form 4 November 2025; Accepted 20 February 2026
Available online 28 February 2026
0968-090X/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

in facilitating efficient travel for residents (Wen et al., 2024). Due to their function of integrating into urban multimodal transport networks through distributed access design, these hubs exhibit complex internal spatial structures, including multi-level architectural layouts and diverse functional areas such as waiting zones, commercial spaces, and other passenger service facilities. Under the interaction of various transport modes and commercial operations, passenger travel purposes, activities, and behaviors within the hub become increasingly heterogeneous. Therefore, effectively guiding and managing pedestrian flows within such complex integrated hub spaces is of vital importance for transport operation and management authorities. It is thus essential to conduct in-depth research on individual pedestrian wayfinding behaviors in hub environments to identify the key factors influencing route choice and uncover the underlying cognitive mechanisms, thereby providing both theoretical foundations and practical guidance for the optimization of spatial layout and the development of intelligent wayfinding systems within transport hubs.

The essence of pedestrian wayfinding behavior lies in the process by which individuals locate themselves within a space and select a route based on the integration of personal needs and spatial attributes. Currently, nearly all major breakthroughs in the field of pedestrian wayfinding behavior have revolved around three key dimensions: (1) methods for collecting and acquiring behavioral data. (2) the quantification of spatial factors influencing wayfinding decisions, along with the characterization of pedestrian heterogeneity. (3) modeling approaches for wayfinding behavior based on multi-source behavioral data. Accordingly, the following review will examine existing research on pedestrian wayfinding from these three perspectives, with a focus on identifying current limitations at both theoretical and technical levels.

1.2. Review of behavioral data collection methods

Behavioral data collection methods can be categorized into field observation and controlled experiments based on virtual reality (VR) environments. Early studies predominantly relied on field-based observation approaches (Asano et al., 2010; Kobes et al., 2010; Yang et al., 2003; Barati et al., 2021; Jin et al., 2024), where in researchers constructed simplified experimental settings (Li et al., 2019; Haghani and Sarvi, 2017a; Srinivasan et al., 2017) or selected specific activity areas within real-world public spaces (Fang et al., 2010; Wong and Lo, 2007; Nilsson and Johansson, 2009). Video surveillance equipment, such as fixed cameras (Wang et al., 2011), was used to record the movement trajectories of either recruited participants or incidental pedestrians during their exit selection process, followed by subsequent analysis and modeling. As these methods are conducted in real physical environments, they are subject to actual crowd density, ambient noise, lighting conditions, and other environmental variables, thereby offering high ecological validity (Feng et al., 2021). However, field-based observation approaches for studying wayfinding behavior suffer from limited experimental controllability due to external interferences, such as unexpected incidents or the disruptive behavior of other pedestrians (Hoogendoorn, 2004). In addition, they often involve substantial temporal and logistical costs (Feng et al., 2021) For integrated transport hubs, which serve as critical public spaces, it is typically infeasible to conduct fully enclosed field experiments that comprehensively capture passengers' movement trajectories across the entire facility. From a technical standpoint, methods based on mobile phone signaling (Calabrese et al., 2011) or GPS tracebacks (Duives et al., 2019) also fall short in accurately capturing fine-grained wayfinding trajectories within indoor environments, given the spatial resolution limitations of such data (e.g., mobile signaling data generally exceeding 10 meters in accuracy). Moreover, these methods are incapable of reconstructing actual cross-level movements within complex indoor settings.

The studies by Yu et al. (2023b) and Chen et al. (2020) have demonstrated that virtual reality (VR) technology provides an effective means for reconstructing the complex activity chains of pedestrians within transport hub environments. However, the majority of existing research has predominantly focused on emergency evacuation scenarios, with relatively limited efforts dedicated to employing VR-based behavioral experiments to retrospectively examine the full travel process of pedestrians under normal operational conditions in integrated transport hubs and to model their decision-making behaviors. Moreover, prior studies have confirmed that differences between immersive and non-immersive VR experimental setups do not exert significant influence on pedestrian wayfinding strategies (Feng et al., 2022b). Non-immersive VR experiments additionally offer advantages such as lower implementation costs and reduced risk of motion sickness. Accordingly, this study constructs a VR platform that closely replicates the operational conditions of an integrated transport hub under normal circumstances and conducts non-immersive behavioral experiments. This approach ensures experimental repeatability while allowing precise reconstruction of participants' complete travel processes via the simulation platform, thereby providing robust input data for accurately modeling pedestrian wayfinding decision-making within hubs.

1.3. Review of influential factors on wayfinding behavior

Influential factors in pedestrian wayfinding behavior can be broadly categorized into environmental and individual factors based on their origin (Haghani and Sarvi, 2018). Existing studies have widely considered environmental variables such as path length, obstacle avoidance, and dimensional differences in spatial configurations (Basu and Sevtsuk, 2022; Andresen et al., 2018; He et al., 2023), establishing significant associations between wayfinding behavior and preferences or cognitive biases related to shortest paths, obstacle occlusion, and the use of local versus global spatial information. With the advancement of virtual reality (VR) technologies, researchers have been able to more intuitively capture visual perception during the wayfinding process through devices such as eye trackers (Chen et al., 2020; Cosma et al., 2016). This has enabled the integration of sensory factors–such as the observation of directional signage (Feng et al., 2021; Yu et al., 2023) and spatial visual field information (Zhou et al., 2021; Liu and Neisch, 2023)–into behavioral analyses. However, eye-tracking-based methods can only record signage that is explicitly fixated upon by the subject, thereby omitting implicit spatial cues such as route visibility and spatial connectivity (van der Hoeven and van Nes, 2014). Furthermore, data derived from eye tracking exhibit limitations in generating structured numerical features suitable for dynamic

discrete choice modeling. Specifically, these methods are constrained in their ability to simultaneously capture and encode multi-scale spatial information, both in terms of breadth and depth, within a three-dimensional environment and to embed such information effectively into sequential behavioral decision models.

Individual factors are primarily reflected in pedestrians' heterogeneous attributes, including endogenous characteristics such as environmental familiarity (Kinateder et al., 2018; Feng et al., 2024), gender (Dulebenets et al., 2019; Lieu and Guhathakurta, 2025; Yin and Cherchi, 2024), and stress levels in emergency scenarios (Yu et al., 2023a). These factors have been extensively validated as significant determinants of route choice behavior during wayfinding processes. However, in the context of high-intensity mixed-use development within hub spaces and the growing deployment of intelligent mobility services, the differentiated wayfinding behaviors associated with heterogeneous travel purposes, such as shopping, ticketing, and waiting during routine hub operations have yet to be thoroughly examined or quantitatively assessed. Furthermore, empirical studies validating the role of two-dimensional (2D) navigation aids in enhancing spatial cognition and supporting passengers' diverse activities remain limited. While numerous studies have conducted wayfinding experiments in both three-dimensional (3D) (Dong et al., 2022; Kwee-Meier et al., 2019; Dubey et al., 2019) and two-dimensional virtual environments (Dong et al., 2021; Tong and Bode, 2024), and have compared the effects of dimensional differences on wayfinding behavior (Dai and Li, 2024), these studies have only preliminarily confirmed that dimensionality is a key factor contributing to spatial cognitive disparities. However, it is the integration of 2D visual interfaces with 3D spatial environments that more accurately reflects real-world travel scenarios, offering a more ecologically valid framework for assessing the effectiveness of navigation assistance tools in complex settings.

1.4. Review of wayfinding behavior modelling approaches

Existing studies have widely employed Logit-based discrete choice models to analyze pedestrian wayfinding behavior (Dai et al., 2024; Jin et al., 2024; Li et al., 2019; Haghani and Sarvi, 2017b; Dong et al., 2020), wherein influencing factors are typically incorporated as static variables or interaction terms to estimate pedestrians' preferences in environments with more than two available exits or pathways. In addition, some researchers have modeled the wayfinding process as an objective function jointly influenced by environmental and individual factors. For instance, Crociani et al. (2016) formulated a route choice model that integrates expected walking time, congestion levels, and nearby agents into a composite utility function. Yue et al. (2014) proposed a hybrid strategy combining distance, time, and cognitive familiarity to model pedestrian route choice in evacuation scenarios. However, recent findings (Alinaghi et al., 2025) suggest that pedestrian spatial cognition is inherently dynamic, evolving throughout the wayfinding process. Consequently, behavioral data and conclusions derived from simplified environments and static modeling frameworks are inadequate for explaining the dynamic and heterogeneous travel behaviors exhibited by pedestrians in complex integrated hub spaces. Furthermore, conventional static Logit-based models are incapable of revealing pedestrians' latent spatial cognition levels. Therefore, there is a critical need for novel modeling frameworks that can jointly capture latent cognitive states and dynamic wayfinding behaviors to more accurately reflect pedestrian decision-making in realistic, high-complexity environments.

1.5. Summary of existing research

The literature review reveals the following findings: (1) In terms of behavioral data collection and acquisition, methods based on field observations, GPS, or mobile phone signaling generally suffer from limited experimental controllability, insufficient spatial resolution, and high implementation costs. Although existing studies have confirmed that VR technology can effectively reconstruct complex activity chains within transport hubs, most applications have focused on emergency evacuation scenarios. Few studies have employed VR to accurately retrospectively capture the complete travel process of pedestrians under normal operating conditions. (2) Regarding spatial determinants of wayfinding decisions and pedestrian heterogeneity, prior research has predominantly emphasized physical spatial attributes–such as route length and obstacle avoidance–and explicit visual cues observed through eye-tracking data, while neglecting implicit visual indicators such as path connectivity and spatial continuity. In addition, the consideration of pedestrian heterogeneity has largely been limited to gender, stress levels, and environmental familiarity. The differentiated travel purposes within hubs, such as shopping or ticketing and the role of navigation aids (e.g., two-dimensional navigation tools) in shaping spatial cognition and behavioral heterogeneity remain underexplored. (3) In the domain of dynamic modeling of wayfinding behavior and cognitive evolution, most existing studies rely on Logit-based frameworks to construct static discrete choice models, which fail to account for the evolving nature of pedestrian spatial cognition and are thus insufficient to uncover latent dynamic cognitive state transitions in complex environments. Although the dynamic nature of spatial cognition has been empirically validated, there is a lack of analytical frameworks capable of integrating latent cognitive states with dynamic modeling of wayfinding behavior in a unified structure.

Building upon the limitations identified in existing research, this study makes the following key contributions:

(1) Unlike traditional research paradigms focused solely on the shortest or fastest evacuation in emergency scenarios, this study constructed a VR scenario of an integrated transportation hub under routine operating conditions and conducted behavioral experiments. The design explicitly incorporates key heterogeneous factors such as service facility selection, commercial area attraction, and use of 2D navigation aids, to examine the differentiated path choice behaviors of participants with varying travel demands. The findings directly support management decisions for optimizing facility configuration, commercial layout, and wayfinding system design in transport hubs.

(2) To achieve a more comprehensive understanding of participants' wayfinding processes, the study integrated space syntax theory, graph embedding techniques, and semantic segmentation methods to extract multiple visual and spatial feature indicators. These include multi-scale visual breadth, high-order semantic attributes of spatial nodes, and visual depth metrics, as well as implicit visual

cues such as field-of-view area and spatial connectivity. This approach produced an interpretable set of explanatory variables that fuse spatial knowledge with visual information, enabling a detailed quantification of decision-making differences among pedestrians with varying travel purposes (e.g., shopping, ticketing) and arrival modes within the hub.

(3) A dynamic wayfinding behavior modeling framework is proposed, combining a multi-state Hidden Markov Model (HMM) with a Logit model, driven by the mechanism of decision cascading effects. By incorporating mutual information decay analysis, the model quantitatively captures the temporal dependence, persistence, and attenuation patterns of historical decisions on current choices. This mechanism informs the precise tuning of hyperparameters and discretization granularity in the proposed DHMM-Logit model, enabling accurate joint modeling of latent cognitive states and sequential decision-making behaviors.

The structure of this paper is as follows. Section 2 describes the non-immersive behavioral experimental methodology based on a virtual reality (VR) scenario simulating an integrated transport hub. Section 3 presents the types of behavioral data collected and the corresponding preprocessing procedures. Section 4 details the proposed dynamic modeling framework for pedestrian route choice behavior. Section 5 discusses the modeling results and their implications. Section 6 concludes with a summary of the study's key contributions, outlines its relevance to the operational management of integrated transport hubs, and provides directions for future research.

2. VR based behavioral experimental design and data collection

This chapter presents the construction process of the virtual scenario, the behavioral experiment design, and the data collection methods used for studying pedestrian route choice within integrated transport hubs. Section 2.1 outlines the development process of the VR scenario, which is driven by real-world spatial data, including the modeling methodology and specifications of the experimental equipment. Section 2.2 introduces the pre-experiment questionnaire designed to capture participants' characteristics, such as demographic attributes and travel habits, and reports the aggregated characteristics of the recruited participant group. Section 2.3 describes the full procedure of conducting the behavioral experiment within the virtual scenario. Finally, Section 2.4 presents the various types of behavioral data collected from participants during the experiment, along with the corresponding data acquisition methods.

2.1. Virtual scenario construction

The study constructs a virtual scene of the integrated transport hub based on a three-layer framework, comprising the physical layer, I/O layer, and virtual layer as illustrated in Fig. 1. Specifically, at the physical layer, the input data include the hub's actual spatial layout (CAD drawings of the station building), the relative coordinates of passenger service facilities (e.g., security checkpoints, ticketing machines, and fare gates), high-speed railway timetables, high-resolution on-site photographs, and ambient audio recordings. These inputs collectively support the reconstruction of realistic architectural textures and soundscapes within the virtual environment. On this basis, a 3D model of the hub was developed using Autodesk Maya 2021, with a scaling ratio of 1unit = 1m and subsequently imported into Unity 2021.3.21f1c1 in FBX format for lossless rendering. Dynamic information, such as train schedules is mapped into the virtual environment as real-time variables via the I/O layer, enabling dynamic updates of relevant scene elements. All interactive logic and user interface prompts are programmed in C# using Visual Studio 2021 under the .NET 4.x framework. For human-computer interaction, Unity's Input Manager is employed to capture keyboard and mouse displacement events, allowing participants to navigate and rotate freely within the virtual environment from a first-person perspective. To ensure smooth desktop operation, the experiment was run on a Dell Inspiron 16 Plus workstation (Windows 10 64-bit OS, Intel i7-12700H 2.30 GHz processor, 32 GB RAM), with a display resolution of 3072 × 1920 pixels and a refresh rate of 90 Hz. In terms of visual perspective and movement settings, a first-person view is implemented using Unity's Cinemachine Camera, with a field of view configured to 110° (Feng et al., 2022a), supporting full 360° horizontal rotation. The default walking speed of the virtual pedestrian is set to 1.5 m/s (Murtagh et al., 2021), replicating realistic locomotion in the virtual layer.

The virtual scenario used for the behavioral experiment was developed based on Hongqiao Station, located in Minhang District, Shanghai, China. This integrated transport hub spans over 1.3 million square meters, with a waiting hall area of approximately 11,340 square meters, and is seamlessly connected to Shanghai Hongqiao International Airport, Shanghai Metro Lines 2, 10, and 17, as well as the city's road. Within this environment, passengers are confronted with multiple spatial route options, resulting in behavioral data that exhibit a high degree of strategic variability and dynamic adjustment, thereby offering a rich dataset for developing more explanatory and nuanced wayfinding models. Fig. 2 presents both real-world views of the Hongqiao Integrated Transport Hub and multi-perspective screenshots of the virtual scene constructed in this study. Owing to the integration of diverse passenger service zones–including waiting areas, ticketing zones, security checkpoints, and commercial areas and the presence of non-intuitive, multi-zonal planar pathways, the spatial layout is characterized by complex partitioning and indirect routing. This setting provides an ideal experimental foundation for modeling the joint influence of individual heterogeneity, visual perception, and spatial knowledge on pedestrian wayfinding decisions.

2.2. Preliminary questionnaire survey

Prior to conducting the formal behavioral experiment within the VR simulation of the transport hub, a pre-experiment questionnaire survey was administered to ensure that the subsequent behavioral trials would accurately reflect the heterogeneous route choice

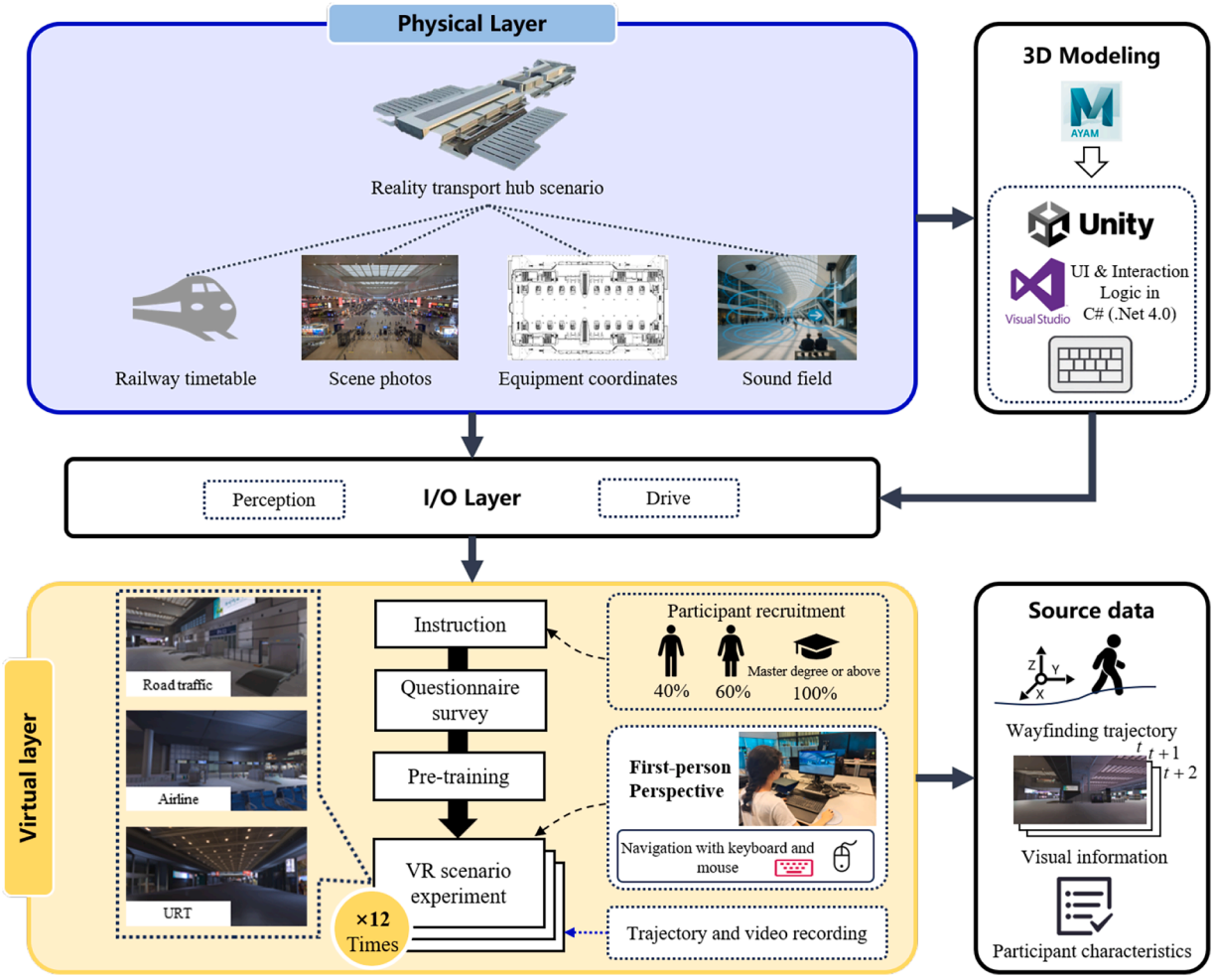


Fig. 1. Virtual scenario construction and behavioural experiment framework for integrated transport hubs.

behaviors exhibited by travelers in integrated transport hubs. This survey systematically collected each participant's individual characteristics and typical travel habits within hub environments. The questionnaire (see Appendix A) covered a range of demographic variables, including gender, age and prior experience with railway travel. In addition, it addressed participants' engagement in non-essential yet common behaviors beyond core procedures such as security screening and waiting, including whether they typically retrieve paper tickets from vending machines, shop in commercial areas during the waiting period, and their average advance arrival time at the station. The gathered information was intended to inform the experimental setup and ensure alignment with real-world behavioral diversity.

In October 2024, the study recruited 60 graduate students (30 male, 30 female; age range: 21-30 years; mean age: 24.3 years) from Beijing Jiaotong University via social media platforms and campus bulletin announcements to participate in the behavioral experiment. Prior to participation, all individuals completed an online questionnaire and provided informed consent by checking a confirmation box. The questionnaire results, summarized in Table 1, reveal significant variation among participants in terms of non-essential travel behaviors. All 60 participants reported prior experience with VR behavioral experiments (60/60, 100%) and no participant reported experiencing motion-sickness symptoms in this study (0/60, 0%) based on the post experiment questionnaire. Notably, advance arrival time at the station exhibited a marked right-skewed distribution: approximately 80% of participants reported arriving 30 to 60 min in advance, while 16% arrived within 30 min, and only 2% arrived more than 2 h early. This distribution shows clear positive skewness (Fisher skewness coefficient $= 0.76 > 0$), indicating a concentration toward moderate early arrivals. These findings provided critical input for designing personalized experimental paths in the subsequent behavioral trials, ensuring that the simulation scenarios reflected realistic and diverse travel behaviors.

2.3. Behavioural experiment procedure

The VR behavioral experiment consisted of three main stages: pre-experiment participant training, the formal wayfinding tasks, and continuous data collection throughout the entire process. Based on the previously collected questionnaire responses, each participant



Fig. 2. Real and virtual environment views of Shanghai Hongqiao integrated transport hub.

Table 1
Descriptive statistics of participant characteristics based on questionnaire survey (consistent with the questionnaire survey content in Appendix A). For the Scenario-Familiarity variable in the table, the study defines participants who have visited the hub scenario at least once in real life as familiar with the scenario.

Items	Variable type	Level	Statistical quantity
Age	Integer variable	–	Average 24.3
Gender	Nominal variable	1: Male	30 (50%)
		0: Female	30 (50%)
Scenario familiarity	Nominal variable	1: Familiar	24 (40%)
		0: Unfamiliar	36 (60%)
Time of early arrival	Categorical variable	1: ≤ 0.5h	10 (17%)
		2: 0.5–1h	44 (73%)
		3: 1h–2h	5 (8%)
		4: ≥ 2h	1 (2%)
Ticket collection	Nominal variable	1: Ticket collection	23 (38%)
		0: No ticket collection	37 (62%)
Ticket collection sequence (for participants who engage in this behaviour)	Nominal variable	1: Get ticket first	23 (100%)
		0: Others	0 (0%)
Shopping	Nominal variable	1: Shopping	21 (35%)
		0: No shopping	39 (65%)
Shopping behaviour sequence (for participants who engage in this behaviour)	Categorical variable	1: Before heading to the waiting area	4 (19%)
		0: After heading to the waiting area	17 (81%)

was assigned a personalized experimental task sequence tailored to their reported travel behavior, enabling accurate reconstruction of their realistic route choice behavior chains within the transport hub environment. During the training phase, participants were introduced to the VR scenario using a non-immersive interface, navigating a virtual space distinct from the main experimental scene. Using a keyboard and mouse to control a first-person perspective, participants practiced three fundamental navigation operations: moving forward, rotating the view, and turning while in motion. Once a participant successfully completed all three operations, they were considered sufficiently trained and proceeded to the formal wayfinding task within the virtual simulation of the transport hub.

During the formal wayfinding phase, each participant was required to complete 12 independent wayfinding tasks within the virtual scenario. Based on their self-reported non-essential behaviors from the pre-experiment questionnaire, the system inserted corresponding task nodes such as ticket collection and shopping-into each participant's route to simulate their actual activity chains. For



(a) Example of a first-person perspective behavioural experiment with 2D floor plan

(b) Example of a first-person behavioural test perspective with no 2D floor plan

Fig. 3. Example comparison of hub plane 2D navigation and general wayfinding first-person participant view.

example, a participant's assigned travel sequence might follow the pattern: "entrance→ticket collection (optional)→waiting→shopping (optional)→ticket inspection→boarding." At the beginning of each task, the system assigned a designated entrance and a ticket inspection gate as the origin and destination, respectively. A task was considered complete when the participant successfully visited all required task nodes and ultimately reached the assigned ticket gate. To explore and compare differences in wayfinding decisions among passengers arriving at the hub via different connecting transport modes–such as metro, air travel, or taxis, which typically link to distinct entrances in the hub–the experiment evenly distributed participants' initial positions across the four available station entrances. Each entrance served as the starting point for three of the 12 tasks, with the order randomized across participants. This design aimed to realistically simulate the diverse arrival patterns associated with multimodal access to the integrated transport hub.

To investigate and compare the influence of spatial navigation aids on passengers' wayfinding strategies and service facility choices within the hub, a subset of participants was randomly provided with a 2D navigation during the wayfinding tasks, as illustrated in Fig. 3. Participants who received the 2D navigation were able to access detailed information about the spatial layout, including the locations of ticket vending machines, fare gates, and various types of shops within commercial areas, as well as their own current position within the environment. The proportion of participants provided with the floor plan was set at 50% of the total sample, thereby enabling the establishment of a control group and a treatment group. This setup supports the quantitative evaluation of the floor plan's effect on participants' route choices and spatial decision-making strategies within the integrated transport hub.

2.4. Source data collection

The trajectory data were captured in the form of three-dimensional spatial coordinates, documenting participants' positional displacements within the virtual scenario. Data were sampled at a frequency of one frame per second, ensuring sufficient spatial resolution to capture detailed movement dynamics. Given that the experimental scenario involved multi-level routes, the use of three-dimensional coordinates rather than two-dimensional planar tracking was essential for accurately representing vertical transitions and inter-floor navigation. During the implementation of the behavioral experiment, the VR platform automatically and continuously recorded two core types of behavioral data for each participant throughout all wayfinding tasks: trajectory data and visual field information. Meanwhile, visual field information was obtained through the platform's built-in screen recording functionality, which continuously logged the first-person perspective of each participant as they navigated the virtual scenario. This recording captured the dynamic visual environment experienced by participants in real time. Such data provide an objective empirical basis for subsequent analyses linking route choice behavior with visual environmental cues, enabling the identification of patterns in spatial perception and attention during the wayfinding process.

The collected trajectory and visual field data were automatically archived and annotated by the experimental platform to ensure traceability and facilitate efficient data processing and analysis. The specific methods for behavioral data processing and temporal sequence alignment will be elaborated in detail in Chapter 3.

3. Data processing and analysis

This chapter outlines the methods used for processing and analyzing the behavioral data obtained from the experiment. Section 3.1 presents the procedure for serializing trajectory data into decision point sequences. Section 3.2 describes the techniques employed to extract and quantify visual field information during wayfinding process. Section 3.3 summarizes the individual attributes, including demographic and behavioral characteristics. Section 3.4 explains the methodology for aligning and integrating the various types of data into a coherent and temporally synchronized analytical dataset.

3.1. Trajectory data

To transform continuous wayfinding trajectories into discrete decision point sequences suitable for route choice behavior analysis, the study first defined a set of explicit spatial boundary conditions. A spatial region classification function was then applied iteratively

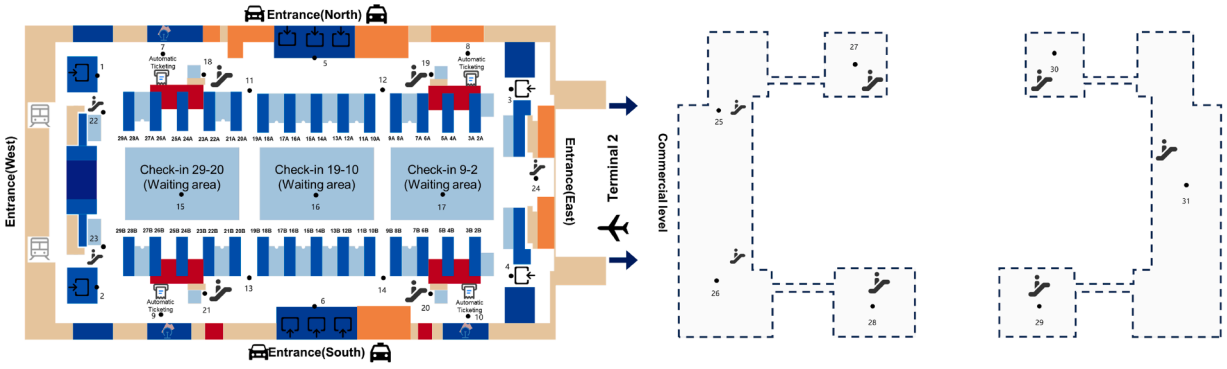


Fig. 4. Example comparison of hub plane 2D navigation and general wayfinding first-person participant view.

to the raw trajectory data to identify and extract the specific decision regions that each participant actually entered during the wayfinding process.

First, the actual spatial layout of the VR-based transport hub was used to identify and label key decision nodes or regions, such as intersections and service facilities (e.g., escalators, ticket vending machines) where participants may face two or more potential route choices. The decision node labeling scheme is illustrated in Fig. 4. We provide node degree distribution data in Appendix B, along with statistics on participant transition frequencies between nodes derived from behavioral experiments. Subsequently, the entire hub space was partitioned into non-overlapping functional regions, each centered on a corresponding decision node. Every region was defined by a polygonal boundary specified by a set of known vertex coordinates. For a given participant trajectory $c_i(t') = \{x_i(t'), y_i(t'), z_i(t')\} \in \mathbb{R}^3, t' = 1, \ldots, T_i$, where T_i is the total number of recorded time steps for trajectory i and t' is the time step. A region membership function $\delta_{i,k_r}(t)$ is used to determine whether the participant is located within the boundary $\mathcal{R} = \{R_{k_r} | k_r = 1, \ldots, M_R\}$ of a functional region k_r at that time step t. When the participant's trajectory enters a region associated with a decision node for the first time, the function records the node index in the participant's decision node sequence. This process progressively forms an ordered sequence of visited decision nodes, which can be formalized as:

$$\delta_{i,k}(t) = \begin{cases} 1, & c_i(t') \in R_k \\ 0, & otherwise \end{cases} \tag{1}$$

$$\begin{aligned} &\mathcal{S} = \{s_1, s_2, \ldots, s_i, \ldots, s_{600}\}, s_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,T_i'}) \\ &\text{st. } t_{\text{in}}(v_{i,1}) < t_{\text{in}}(v_{i,2}) < \cdots < t_{\text{in}}(v_{i,T_i'}), i \in [1, 600] \end{aligned} \tag{2}$$

where v denotes the index of the decision node identified based on the trajectory data. s_i represents the decision sequence generated for the i^{th} trajectory and T_i' is the length of the sequence, i.e., the total number of decision nodes visited. The corresponding time point at which a participant enters a decision node region is denoted by $t_{\text{in}}(\cdot)$ indicating the timestamp of entry into node. For participants assigned non-essential tasks such as ticket collection or shopping, route decisions involving such behaviors require active identification of relevant service facilities within the scene. These decisions may be aided by either visual observation of the environment or by reference to the floor plan (if provided). Upon reaching the respective functional area, participants were instructed to report the completion of the non-essential activity to the experiment facilitator, enabling accurate logging of the node index associated with the service zone. Ultimately, the raw trajectory data for each participant were transformed into a temporally ordered sequence of decision nodes, constrained by the actual entry times into corresponding spatial regions. The result is a comprehensive set of decision sequences $\mathcal{S}$.

3.2. Visual information

The processing of visual field information involves the extraction and computation of both visual depth and visual breadth features. Our previous research has confirmed that variations in the proportion of visible elements, such as routes and obstacles within a pedestrian's field of view significantly influence route choice behavior (Dai et al., 2024). However, approaches that consider only visual breadth lack the capacity to represent the semantic structure of space and often overlook the continuity of spatial depth. For example, two paths may occupy the same proportion of a pedestrian's view, yet one may lead to a navigable corridor while the other is blocked by a wall. In such cases, proportion-based metrics fail to distinguish between these alternatives, resulting in inaccurate predictions of preferences such as the tendency to choose visually continuous corridors. Therefore, in addition to the proportional visibility of spatial elements (i.e., visual breadth), it is essential to incorporate spatial syntactic measures, such as connectivity and integration, as supplementary variables. By integrating knowledge of both visual depth and visual breadth, a more comprehensive understanding of the cognitive processes underlying pedestrian wayfinding behavior can be achieved.

3.2.1. Visual depth information

The computation of visual breadth data was conducted by importing the CAD-based spatial layout of the transport hub into Depthmap (version 10.14.00b). The spatial plane was discretized into a 2m × 2m grid, allowing for the construction of a visibility graph across the entire environment. Based on this graph, key space syntax metrics, such as connectivity, integration, and visibility area were calculated for the locations corresponding to each decision node, providing a quantitative representation of visual accessibility and spatial prominence within the hub.

1) Connectivity I_1 (Range: $[0, +\infty)$) : The connectivity of a target node in the visibility graph is defined as the total number of nodes that are directly intervisible with it. This metric reflects the degree of direct visual connectivity between a given spatial location and its surrounding areas, indicating how visually accessible a region is within the overall spatial configuration.

2) Isovist Area I_2 (Range: $[0, +\infty)$) : The visual area refers to the total spatial area visible from the target location, centered on their point of view. It quantifies the extent of the visual field within the environment, representing how much of the surrounding space is perceivable at a given position.

3) Isovist Perimeter I_3 (Range: $[0, +\infty)$) : The isovist perimeter refers to the perimeter length of the visible boundary from the pedestrian's target location. It captures the geometric complexity of the visual field, with larger or more irregular perimeters indicating more intricate or fragmented views from the observation point.

4) Visual Clustering Coefficient I_4 (Range: $[0, 1]$) : The visual integration metric describes the compactness of the spatial layout and the degree of spatial clustering within the visual field. It reflects how integrated or segregated a location is in relation to the overall space, indicating the ease with which other areas can be visually accessed from a given point.

$$I_4(v) = \frac{2 \cdot e_c}{k_c(k_c - 1)} \tag{3}$$

where e_c denotes the number of actually existing visible connections among the set of visually adjacent nodes $NC(v)$ of node v, while k_c represents the number of visual neighbors of node v.

5) Visual Integration I_5 (Range: $[0, +\infty)$) : The global visual integration value is defined as the reciprocal of the average visual steps required to reach all other nodes in the space from the target node. It quantifies the visual accessibility of a target location within the global spatial configuration, reflecting how easily the area can be recognized and reached based on the spatial structure.

$$I_5(v) = \frac{1}{MD(v)} \tag{4}$$

$$MD(v) = \frac{\sum_{u \in V} d(v, u)}{|V| - 1} \tag{5}$$

where $MD(v)$ denote the average visual depth from a target grid v to all other visually accessible points u. $d(v, u)$ represents the shortest visual step length (i.e., the minimum number of steps in the visibility graph) from node v to u. $|V|$ is the total number of decision nodes. Fig. 5(a)–(e) respectively illustrate the spatial distributions of the five space syntax metrics, including connectivity, isovist area, isovist perimeter, visual clustering coefficient, and visual integration computed across all discretized grid cells within the environment. Since the value ranges of these five metrics differ significantly, Min-Max normalization was applied prior to visualization to scale all values uniformly to the range $[0, 1]$.

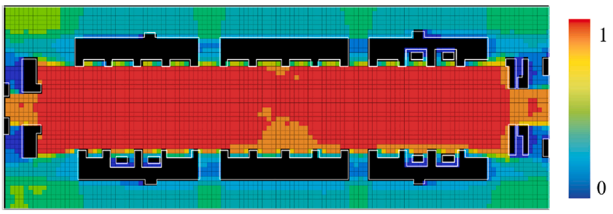
3.2.2. Visual breadth information

For the extraction of visual breadth information, the VR experimental platform's built-in real-time screen recording functionality was used to capture each participant's first-person view images at different decision nodes. These images were recorded at a resolution of 1920 × 1080 pixels in RGB format. To quantify participants' visual attention to different types of elements within their field of view, the study employed the state-of-the-art semantic segmentation network SCTNet (Xu et al., 2024). This model was applied to the recorded view images to perform semantic segmentation, enabling precise identification and classification of key spatial elements. Following segmentation, the proportional area occupied by each semantic category was computed within each image, thereby generating a set of visual breadth features corresponding to each decision node.

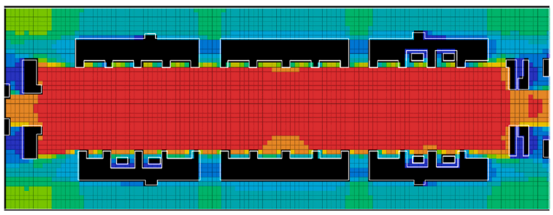
Specifically, we selected 200 field-of-view images and used Labelme to segment and annotate paths, obstacles, and spatial areas within the images, which served as the training set for the semantic segmentation network SCTNet. The remaining images were used as the test set. Fig. 6 illustrates the architecture of the semantic segmentation model employed. We adopted the same network parameters as the original SCTNet and achieved an average pixel classification accuracy of 94.75% after 200 epochs of training. By using the trained model to assign predefined class labels to pixels in the images, we were able to detect and extract the shapes of various elements within the field of view, and further calculate the proportion of each element in the image. Fig. 7 presents an example of radar charts showing the proportion of spatial elements in the field of view at different decision nodes within the hub.

3.3. Participant personal attributes

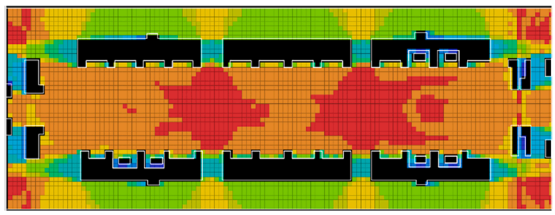
Participants' personal attribute data were obtained through a detailed pre-experiment questionnaire. To avoid the reduction of statistical power caused by an insufficient sample size resulting from an excessive number of independent variables, the study conducted sample size adequacy and statistical robustness tests (see Appendix C). Using clustered bootstrap, randomization permutation testing, and Monte Carlo power analysis, we addressed three key issues: (1) variance underestimation due to within-participant correlation across multiple activity chains, (2) exclusion of behavioral effects arising from random participant-level variability, and (3)



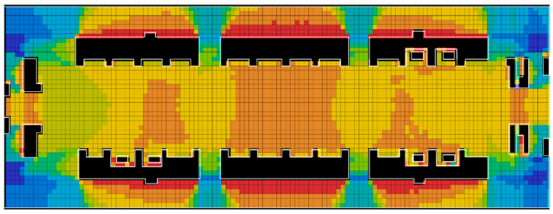
(a) Connectivity visualization diagram



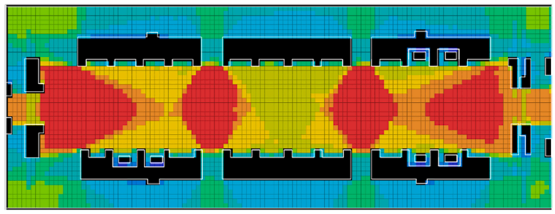
(b) Isovist area visualization diagram



(c) Isovist perimeter visualization diagram



(d) Clustering coefficient visualization diagram



(e) Visual integration visualization diagram

Fig. 5. Visualization the distribution of five spatial syntax indicators for all discrete grids within the hub space. To simplify the calculation process, the study ignored objects such as seats and security equipment that are low in height and have little impact on visibility when creating the visualization diagram. The scale of each discrete grid is 2m×2m.

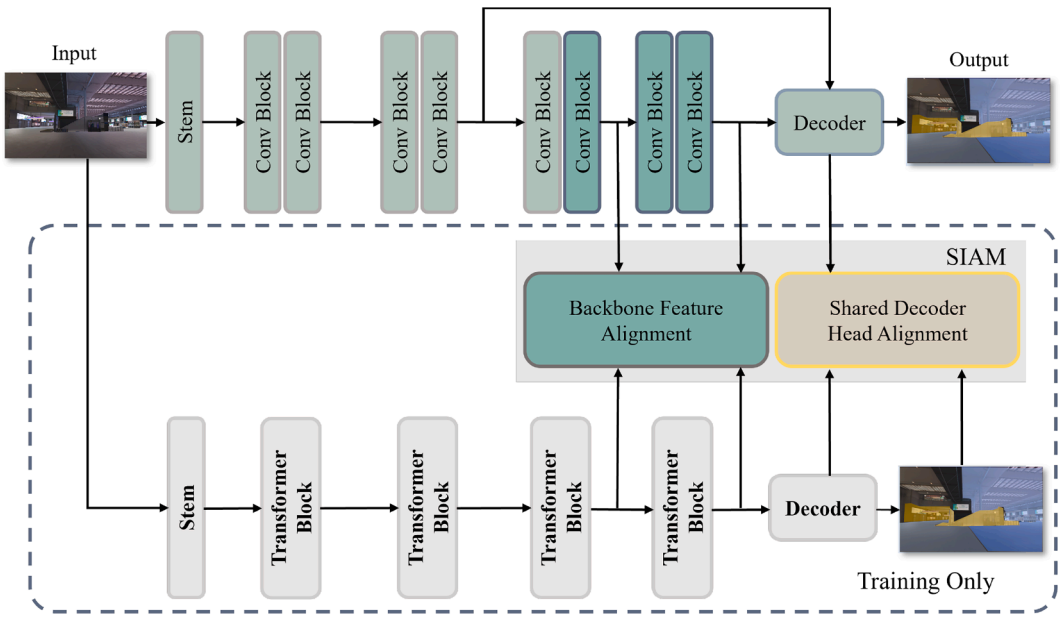


Fig. 6. The architecture of SCTNet. The model follows a single-branch design: only a CNN backbone is employed at inference time, while an additional Transformer-based semantic branch is attached in parallel during training. The two branches are aligned via a Semantic Information Alignment Module (SIAM), thereby combining the efficiency of convolutions with the global semantic awareness of Transformers. The network parameters of the model studied here are identical to those used in the model proposed by Xu et al. (2024).

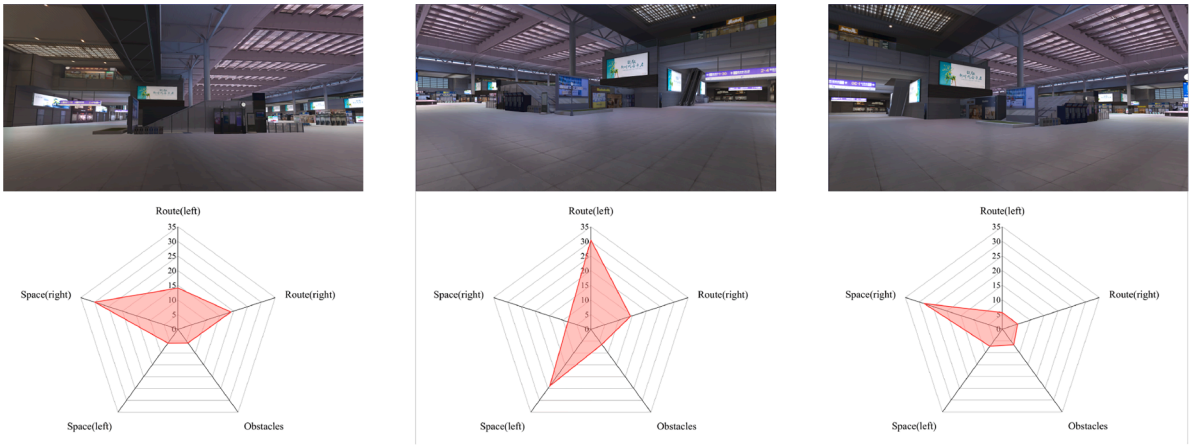


Fig. 7. The radar charts show the proportion of pedestrian vision occupied by five different spatial elements. We extract walkable routes, obstacles, and the space above the routes in the pedestrian's field of view, and calculate their proportion in the field of view through semantic segmentation.

the detection of main effects of heterogeneity factors under the current sample size. The randomization permutation test yielded a two-sided p-value of 0.0045 (< 0.05), indicating sufficient evidence for a true nonzero effect under the current sample. The 95% confidence interval derived from the clustered bootstrap was [97.26, 323.83], which does not cross zero, demonstrating that the treatment effect remains significant even after accounting for within-participant correlation. The Monte Carlo power analysis further showed that, under the current sample size, the experiment would detect the same main effect as significant with approximately 85.2% probability at $\alpha = 0.05$, confirming that the current sample provides adequate statistical power for effect detection.

Meanwhile, the significance tests of heterogeneity attributes on behavioral outcomes (see Appendix C) indicate that gender has no significant effect ($p > 0.05$) and is therefore excluded as an independent variable in subsequent analyses. Consequently, the final set of personal attributes included in the analysis comprises three indicators: early arrival time, familiarity with the scenario and usage of a 2D floor map.

3.4. Data alignment and serialization

To integrate trajectory data, visual information, and participants' personal attributes into a unified input format for subsequent wayfinding behavior modeling, this section outlines the process of aligning and serializing the multi-type behavioral data obtained in the previous sections. Using the decision sequence sets (as established in Section 3.1) as the foundation, we constructed a sequential data structure for each participant across all tasks. This structure incorporates the ID of each decision node visited, the travel purpose (a categorical variable derived from the predefined task flow, including categories such as waiting, shopping, ticket collection, and ticket check), the task completion status (a binary variable indicating whether the participant reached the facility or area corresponding to the current travel purpose; a value of 1 denotes completion for example, arriving at the designated ticket gate under the 'waiting' purpose), the ID of the next target node in the subsequent time step, and the Euclidean distance between nodes. In this study, we define two types of distance metrics. The first is local distance, which refers to the Euclidean distance between the participant's current node and the next node after making a decision at the current step. The second is global distance, defined as the shortest distance from the current node to the node (area or facility) associated with the next travel purpose. We illustrate these two distance concepts in Fig. 8. These metrics are designed to capture whether pedestrians exhibit planning behavior during the wayfinding process within the spatial environment.

Meanwhile, two types of visual information: visual breadth indicators (obtained in Section 3.2.2) and visual depth information (obtained in Section 3.2.1) were aligned respectively with the current and next decision nodes. Specifically, the visual breadth data were associated with the current decision node, while the visual depth data were aligned with the next node to be reached. This mapping approach is based on the rationale that during wayfinding, pedestrians can only directly perceive the visual breadth of their current location. Their willingness to proceed toward a candidate node is often driven by visual cues or spatial knowledge suggesting that the direction ahead is more accessible and coherent in terms of visual depth. By assigning breadth information to the current node and depth information to the subsequent node, the behavioral model at each time step is provided with both perceptual information (what is currently seen) and expectational information (what is anticipated at the next step), thereby constructing a complete sequence of visual features. The validity of this alignment strategy will be further evaluated in Chapter 5.

Finally, participants' personal attribute data (as defined in Section 3.3) were matched using unique participant identifiers to the serialized behavioral data described above. This ensured strict correspondence and integration across all data types at the decision-point level. The resulting dataset is a comprehensive and structured multidimensional sequential dataset, as illustrated in Table 2. The integrated dataset clearly captures both the individual heterogeneity of each participant and their dynamic decision-making processes throughout the experiment.

Table 2
Examples of pedestrian wayfinding behaviour data constructed in this study. ID refers to the participant's assigned number during the experiment, and *Exp_num* indicates the round of the wayfinding experiment (ranging from 1 to 12). *Step* denotes the decision-making time step, which differs from the time step used in the trajectory sequence in Section 3.1. In this study, the behavioral sequence is defined as event-driven: a new time step is recorded whenever a participant departs from one decision node and arrives at the next. L_t represents the ID of the decision node where the participant is currently located at that step. M_s indicates the participant's current travel purpose at this step (ticket collection; 1: shopping; 2: waiting; 3: return to the waiting area after shopping; 4). D_t is a binary variable that reflects whether the travel purpose at the current location has been fulfilled. L_d is the Euclidean distance from the current node to the next decision node (Next location), which is referred to in the study as local distance. G_d represents the shortest distance from the current node to the node where the next travel purpose is to be completed. This is computed by constructing an adjacency matrix between nodes and applying Dijkstra's shortest path algorithm, where edges are established only between directly accessible nodes and weighted by their Euclidean distances. V_v, V_o and V_s denote the proportions of visible route to the next node (Next location), overhead space, and obstacles, respectively, within the participant's field of view at the current step. $I_1 - I_5$ represent five spatial syntax indicators. Participant heterogeneity attributes-including whether the participant used a 2D layout map during the wayfinding task, whether they were familiar with the environment, and their early arrival time-are represented by H_t, H_s and H_y, respectively.

ID	Exp_num	Step	L_t	M_s	D_t	L_d	G_d	V_x	V_y	V_o	I_1	I_2	I_3	I_4	I_5	L_{e+1}	H_t	H_s	H_y
1	1	1	2	3	0	75	195	11.66	36.40	5.46	223	9030	825	0.87	19.3	15	1	0	2
1	1	2	15	3	0	135	135	24.99	25.16	3.07	216	8739	826	0.92	15.0	16	1	0	2
1	1	3	16	3	0	105	105	20.71	32.11	1.14	223	8973	834	0.87	19.4	17	1	0	2
1	1	4	17	3	1	90	93	44.27	21.15	2.91	216	8701	758	0.85	18.7	24	1	0	2
1	1	5	24	3	0	22	90	43.17	21.36	3.71	216	8701	758	0.86	18.8	31	1	0	2
1	1	6	31	4	0	22	93	45.57	24.91	9.35	216	8701	758	0.86	18.8	24	1	0	2
1	1	7	24	4	0	90	194	43.17	21.36	3.71	223	8973	834	0.87	19.4	17	1	0	2
1	1	8	17	4	1	105	105	44.27	21.15	2.91	216	8739	826	0.92	15.0	16	1	0	2
2	1	1	2	1	0	73	191	19.05	30.02	5.77	223	9030	825	0.87	19.3	15	1	0	2
2	1	2	15	1	1	134	134	24.99	25.16	3.07	216	8739	826	0.92	15.0	16	1	0	2
									… … …										
12	10	1	6	2	0	50	169	29.57	21.91	30.35	268	4002	693	0.60	11.9	14	0	3	3
12	10	2	14	2	0	68	126	7.25	30.05	8.23	223	8973	834	0.87	19.4	17	0	3	3
									… … …										

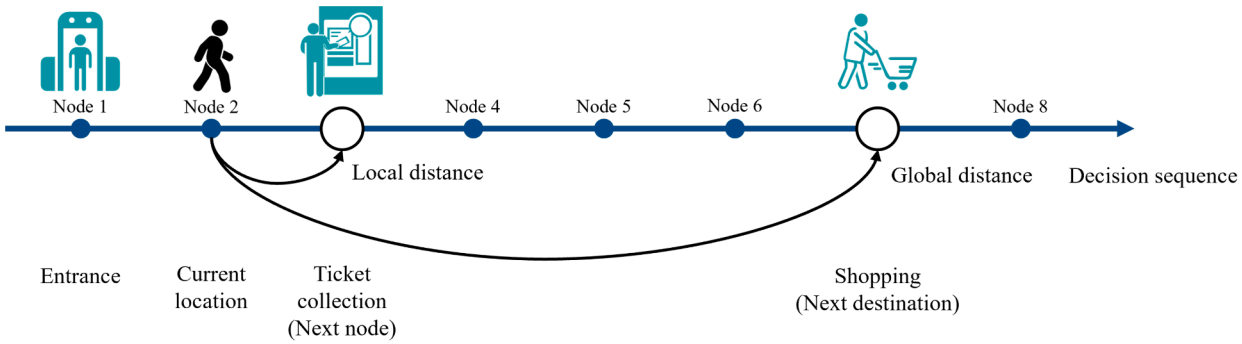


Fig. 8. Illustration of the two distance concepts defined in the study. Local distance refers to the distance from the pedestrian's current location to the next node chosen for movement. Global distance refers to the distance from the current location to the facility or node corresponding to the next destination (which is 'Shopping' in the diagram, while the current purpose is 'Ticket collection').

4. Methodology

To investigate the mechanisms underlying heterogeneous pedestrian wayfinding behaviors in integrated transport hub scenario, this study proposes the DHMM-Logit (Dynamic Hidden Markov Model Logit) framework. Utilizing the serialized behavioral data collected through preceding virtual reality (VR) experiments, the model captures the dynamic and heterogeneous decision-making processes of pedestrians within such hubs. The overall research framework is illustrated in Fig. 9. Specifically, the DHMM-Logit architecture comprises three key components:

(1) Decision sequence dependency modeling and hyperparameter calibration driven by decision cascades. This component leverages behavioral decision cascading effects to model temporal dependencies between consecutive decisions. Metrics such as Mutual Information (MI) are employed to quantitatively analyze the influence strength and decay mechanisms across decision steps, thereby determining the persistence of historical decisions in the behavioral process.

(2) Latent state identification and segmentation. A multi-state Hidden Markov Model (HMM) is applied to identify latent cognitive states that underlie pedestrian decisions. These states are interpretable and behaviorally meaningful, enabling a fine-grained depiction of how pedestrians transition across distinct cognitive states during the wayfinding process.

(3) Dynamic parameter estimation of multi-factor influences. A dynamic Logit model is integrated with the state transition process and estimated via the Expectation-Maximization (EM) algorithm. This allows for the joint modeling of parameter variations across both state and time dimensions, providing predictive probabilities for selecting specific paths or facilities at each decision point. By jointly incorporating state heterogeneity, parameter dynamics, and decision dependencies, the DHMM-Logit framework significantly enhances the ability to model complex and heterogeneous pedestrian behaviors and spatial cognition strategies within integrated hub environments. Compared with traditional single-structure models, this framework offers superior theoretical interpretability and practical applicability.

In this chapter, Section 4.1 introduces the feature engineering and data preparation methods for vectorizing decision node information and encoding spatial attributes based on behavioral data. Section 4.2 presents the interaction-effect validation and selection approach using tree-based models. Section 4.3 defines the concept of decision cascading and details the computation of related metrics. Section 4.4 provides a comprehensive description of the proposed DHMM-Logit model, including its mathematical formulation and computational procedures.

4.1. Feature engineering data preparation

The purpose of feature engineering is to transform raw wayfinding data into features that better capture the underlying behavioral mechanisms, thereby improving the model's predictive accuracy, especially for unobservable factors such as pedestrians' preferences for spatial information. In this study, the feature engineering process consists of two main components: numerical normalization and spatial feature encoding of decision nodes.

4.1.1. Data normalization

The numerical normalization step ensures that features with different units or scales are comparable during model training, allowing the DHMM-Logit model to assign weights more equitably across variables and improving training stability. In this study, Z-score normalization is applied, which can be formalized as follows:

$$X'_{num} = \frac{X_{num} - \mu_{num}}{\sigma_{num}} \tag{6}$$

where $X_{num} = \{L_d, G_d, V_v, V_s, V_o, I_1, I_2, I_3, I_5, H_y\}$ denotes the original numerical feature matrix (since the raw values of I_4, H_t, H_s already fall within a reasonable range $[0, 1]$, no additional normalization is applied to it). μ_{num}, σ_{num} represent the vectors of means

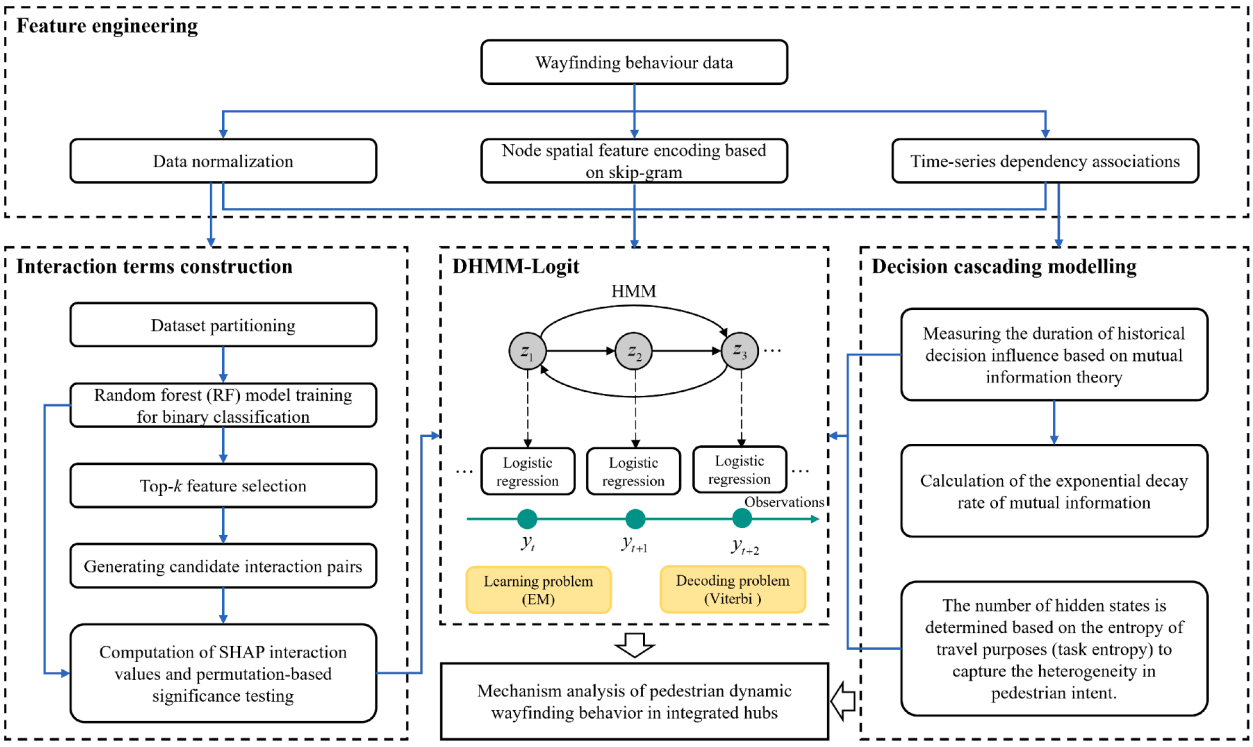


Fig. 9. The methodological framework of this study.

and standard deviations for each feature, respectively. X'_{num} is the resulting standardized numerical feature matrix after applying Z-score normalization.

4.1.2. Spatial feature encoding of nodes

Although participants do not interact with each other during the wayfinding tasks, their individual behaviors collectively reveal structural characteristics and spatial semantics of the hub environment. Simply put, each participant's route choices and the locations they pass through are not random, but rather the result of interactions between environmental layout factors (e.g., route length, obstacles) and task goals (Kessler et al., 2024). Therefore, even in the absence of direct communication, the aggregation of these navigation paths effectively maps out the underlying spatial structure and inter-node relationships within the environment. In this study, we employ a semantic embedding model to analyze the decision sequences of all participants (as obtained in Section 3.1). By learning the co-occurrence patterns between nodes within these sequences, we embed the functional similarity and spatial correlation between nodes into high-dimensional feature vectors. These node embeddings are then used as input features for the DHMM-Logit model, thereby enhancing its accuracy in predicting pedestrians' next-step decisions.

Specifically, based on the set of decision sequences S, we adopt a modified Skip-gram model to train node embeddings. The objective is to maximize the co-occurrence probability between the target node and its context nodes within a sampling window by optimizing a defined loss function $\mathcal{L}(\theta) = -\sum_{v_c \in V} \sum_{u \in W(v_c)} \log p(u|v_c, \theta)$, using a backpropagation mechanism. The model learns spatial feature representations of nodes through this process. In the loss function, θ denotes the model parameters and $p(u|v_c, \theta)$ is the probability of observing context node u given center node v_c. Considering that passenger travel processes within transport hubs follow a certain degree of sequential order, the Skip-gram model adopted in this study introduces a direction-sensitive context window to account for the unidirectional temporal dependencies between nodes–unlike the traditional Skip-gram model (Perozzi et al., 2014) Therefore, during the sliding window sampling process, only the subsequent nodes following the current center node are considered as context, as illustrated in Fig. 10. The report on model hyperparameter values and their impact on DHMM-Logit accuracy under different hyperparameter settings is presented in Appendix D.

Ultimately, we obtain the node spatial feature embedding vectors . Incorporating these spatial embeddings as part of the input to the DHMM-Logit model offers a key advantage: it leverages the collective intelligence embedded in group behavior. By aggregating independent wayfinding decision sequences from a large number of participants, the model captures high-level, group-level decision patterns and latent spatial semantics, expressing them in a compact, low-dimensional form. These embeddings are then used as input corresponding to the pedestrian's most recent spatial decision, enhancing the model's accuracy in predicting sequential decision-making behaviors. Additionally, representing each node with a low-dimensional dense embedding–as opposed to a sparse one-hot encoding significantly reduces both storage requirements and computational costs.

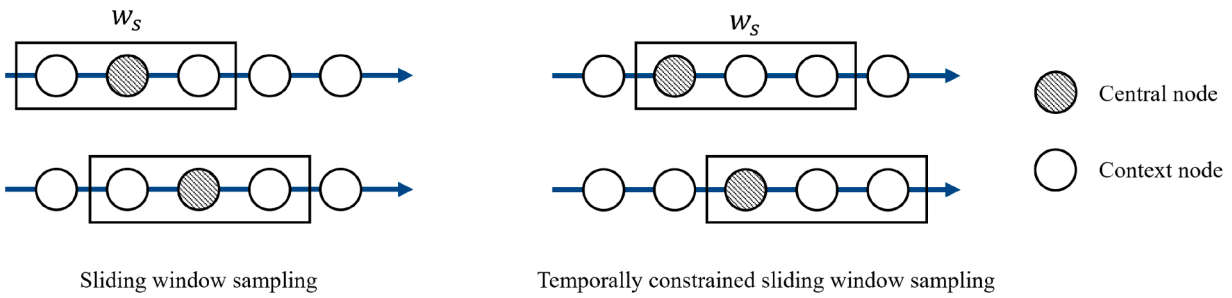


Fig. 10. Comparison between traditional sliding window sampling (left) and temporally constrained sampling approach (right).

4.2. Interaction intensity measurement

In wayfinding tasks, visual information, spatial geometric indicators, and individual attributes often do not act independently but instead exhibit complex synergistic or suppressive interaction effects. To identify and capture these potential interactions among explanatory variables and enhance the model's fit and predictive accuracy on real behavioral data, this study proposes an interaction strength permutation test method based on Random Forests (RF) and SHAP (Shapley additive explanations) values. This approach extracts interaction terms that have a statistically significant influence on pedestrian wayfinding behavior. Leveraging the non-parametric nature of tree-based models, this method effectively handles nonlinear relationships among variables in dynamic wayfinding decisions without requiring prior specification of interaction structures or assumptions of linearity. The true interaction effects are captured through a permutation-based testing mechanism. The steps are as follows:

Step 1: Dataset construction. For each decision node k_p, an independent binary classification Random Forest model is constructed. The dependent variable $R(k_p)$ is defined as follows: if at the current time step the participant is located at a node that is an adjacent node of node k_p in the spatial graph $G = (V, E)$, and the participant makes the next-step decision $L_{t+1} = k_p$ based on the explanatory variables, then we set $R(k_p) = 1$. The graph $G = (V, E)$ is constructed as an adjacency matrix based on the spatial layout of nodes within the environment. The dataset is then split into a training set and a test set in a 7:3 ratio.

Step 2: Main effect feature selection based on RF. A Random Forest model is first trained to identify the most important main-effect features. Using SHAP values, we quantify the contribution of each individual explanatory variable to the model's prediction. The average absolute SHAP value for each feature is computed, and features are then ranked accordingly. To reduce the computational burden of subsequent interaction effect testing, we select the top-k ($k = 5$) features based on their average SHAP importance.

$$\hat{R}(k_p) = \frac{1}{B}\sum_{b=1}^{B} f_b(X_{tr}) \tag{7}$$

in the formula B denotes the total number of trees in the Random Forest model, and $f_b(X_{tr})$ represents the prediction function of the b^{th} tree.

Step 3: Generation of candidate interaction pairs. Pairwise interactions are constructed from the top-k selected features to constrain the scope of interaction term testing, thereby improving computational efficiency.

Step 4: Interaction strength computation. SHAP interaction values ϕ_{α_i,α_j} are employed to quantify the interaction strength of each pair of explanatory variables (α_i, α_j) with respect to the predictive outcomes across different spatial regions. The process can be formulated as follows:

$$\phi_{\alpha_i,\alpha_j} = \sum_{S\subseteq X'\setminus\{\alpha_i,\alpha_j\}} \frac{|S|!(|X'|-|S|-2)!}{(|X'|-1)!}\left[\nu(S\cup\{\alpha_i,\alpha_j\}) - \nu(S\cup\{\alpha_i\}) - \nu(S\cup\{\alpha_j\}) + \nu(S)\right] \tag{8}$$

where X' denotes the selected set of explanatory variables, and S represents a feature subset. The function $\upsilon(\cdot)$ quantifies the contribution of the subset to the model's predictive output, measuring the difference between the model's prediction using only the features in S and the baseline prediction.

Step 5: The statistical significance of interaction terms is assessed via permutation testing to control the false positive rate. Specifically, the training set is subjected to B_r rounds of random permutation to simulate the null distribution of interaction effects. For each permuted label set, a random forest model is retrained, and the corresponding SHAP interaction values are computed to obtain the permuted interaction strength s_p. By comparing these with the previously computed observed interaction strength $s_{obs} = |\phi_{\alpha_i,\alpha_j}|$, the significance of each interaction term is quantitatively evaluated.

$$p_{val} = \frac{\sum_{b=1}^{B_r} I(s_p^b \geq s_{obs}) + 1}{B_r + 1} \tag{9}$$

Through the aforementioned procedure, the study effectively identifies and extracts interaction terms that exhibit a statistically significant influence on wayfinding behavior. These interaction terms are subsequently incorporated as explanatory variables into the DHMM-Logit model.

4.3. Behavioural decision cascading modelling

The hyperparameter calibration method based on decision cascading serves as a critical component in both tuning the hyperparameters of the DHMM-Logit wayfinding model and quantifying the temporal dependencies inherent in dynamic behaviors. Traditional behavioral decision models typically determine hyperparameters, such as the number of hidden states and model order–based on empirical rules or optimal predictive performance criteria (Yu et al., 2023a). However, such approaches often fail to accurately capture the intrinsic dynamic dependency structures of traveler decision sequences across diverse scenarios. To address this limitation, the present study proposes a data-driven method that automatically infers appropriate model hyperparameters by mining the intrinsic information association between decisions at different temporal steps. This approach enables a more faithful representation of participants' actual cognitive and decision-making processes in wayfinding tasks.

We begin by defining the concept of decision cascading as employed in this study. Decision cascading refers to the phenomenon wherein current decisions are continuously influenced by preceding decisions within a travel goal-driven temporal framework, with this influence diminishing over time. To quantitatively characterize this phenomenon, the study introduces the following three key indicators:

Influence duration ℓ_d. This indicator is designed to measure the maximum number of preceding steps over which the current decision at time t exhibits dependence on past decisions, thereby serving as a foundation for determining the appropriate order of the Hidden Markov Model (HMM). The study employs mutual information theory (Shannon, 1948) to quantify the extent to which a decision made at time step $t-\ell_d$ reduces the uncertainty associated with the current decision L_t, thus providing a rigorous quantification of temporal dependency within the decision-making process. The calculation process is as follows:

$$\tilde{I}(L_t, L_{t-\ell_d}) = \frac{\sum_{L_{t-\ell_d}} \sum_{L_t} p(L_{t-\ell_d}, L_t) \log\left(\frac{p(L_{t-\ell_d}, L_t)}{p(L_{t-\ell_d})p(L_t)}\right)}{-\sum_{L_{t-\ell_d}} p(L_{t-\ell_d}) \log p(L_{t-\ell_d})} \tag{10}$$

here, $p(L_{t-\ell_d}, L_t)$ denotes the joint probability distribution of the decision pair $(L_{t-\ell_d}, L_t)$, estimated from the empirical frequency of decision sequences. $p(L_{t-\ell_d}), p(L_t)$ represents the marginal distributions of the respective decision variables. To enable comparison across different lag lengths, the mutual information is normalized by the entropy of the historical decision $L_{t-\ell_d}$, yielding a relative measure of information contribution. In this study, the influence of historical lag length is examined from two perspectives:

(1) It is assumed that when the normalized mutual information $\tilde{I}$ decreases to half or less of its initial value, the influence of past decisions on the current decision is no longer considered a primary contributing factor. The corresponding lag length is defined as the maximum significant influence length ℓ_d^*, representing the temporal half-life of decision influence. This indicator serves to determine the appropriate order of the HMM, striking a balance between capturing temporal dependencies in historical decisions and controlling model complexity.

(2) When the normalized mutual information $\tilde{I}$ first drops below a predefined threshold τ_a, the influence of past decisions on the current decision is considered to have effectively dissipated. The corresponding lag length is defined as the maximum influence duration ℓ_m^*. This indicator is subsequently used to automatically determine the number of temporal bins n_{bin} in the DHMM-Logit model, thereby enabling the discretized learning of continuous decision behaviors and enhancing the model's stability and robustness.

$$\ell_m^* = \min\{\ell_m : \tilde{I}(L_t, L_{t-\ell_m}) < \tau_d \tag{11}$$

$$n_{bin} = \max(2, \lceil T/\ell_m^* \rceil) \tag{12}$$

Temporal decay rate γ_{ℓ_d}. Given that mutual information generally exhibits an exponential decay trend with increasing decision lag (Shen, 2019), the decay rate is defined as the rate at which mutual information diminishes over time. It is computed as follows:

$$\tilde{I}(\ell_d) = exp(-\gamma_{\ell_d} \cdot \ell_d) \tag{13}$$

This indicator is subsequently employed to extend the Hidden Markov chain to incorporate higher-order dependencies and to empirically validate the exponential decay trend of mutual information. Moreover, it facilitates the evaluation of differences in the strength of decision cascading effects across travelers with varying modes of transportation and distinct personal attributes, enabling a deeper understanding of heterogeneity in wayfinding behavior.

Task entropy H_e. This metric is defined as the entropy of the probability distribution over different types of travel tasks (e.g., ticket collection, shopping, waiting for railway train), and is used to quantify the degree of heterogeneity in passengers' travel purposes.

$$H_e = -\sum_{k_e=1}^{K} p(m = k_e) \log_2 p(m = k_e) \tag{14}$$

where, $p(m = k_e)$ denotes the occurrence probability of the kth type of travel purpose, and K represents the total number of purpose categories (set to 4 in this study). A higher entropy value indicates greater heterogeneity in travel purposes within the hub, reflecting increased complexity in pedestrian behavioral patterns. This implies the need for a larger number of hidden states in the Hidden Markov Model to adequately capture the diverse decision-making strategies of various passenger types. In this study, a threshold for

entropy is defined to determine the appropriate number of hidden states M_h

$$M_h = \begin{cases} 2, H_e < 0.5 \\ 3, 0.5 \le H_e < 1.0 \\ 4, H_e \ge 1.0 \end{cases} \tag{15}$$

4.4. Dynamic hidden Markov model logit

During the wayfinding process, a participant's navigation strategy is determined by their latent cognitive perception of the environment, which is inherently unobservable. The Hidden Markov Model (HMM) is well-suited for modeling such latent cognitive states as hidden states. Given that these cognitive states evolve dynamically in response to incoming spatial information and individual heterogeneity throughout the wayfinding process, this study proposes the Dynamic Hidden Markov Model-Logit (DHMM-Logit) framework to capture the influence of spatiotemporal information dynamics, personal heterogeneity, and the transitions of latent cognitive states on navigation decisions. The overall modeling framework consists of three main components: (1) hidden state segmentation, (2) state transition modeling, and (3) dynamic parameter adjustment.

First, hidden states are employed to represent the latent cognitive modes of participants during the route choice decision-making process. The number of hidden states M_h is determined based on the task entropy H_e derived from the decision cascading analysis (Section 4.3). The set of hidden states is denoted as $Z_h = \{z_1, z_2, \dots, z_{M_h}\}$, where each state $z_k, k \le M_h$ corresponds to a distinct underlying cognitive pattern of the participant. The dynamic transitions between these cognitive states are modeled via a state transition process. Incorporating the maximum cascading length indicator, an ℓ_d^*-order Hidden Markov Model $\phi_m = \langle A_m, B_m, \pi \rangle$ is adopted to capture higher-order temporal dependencies in the decision sequences. The model defines initial state distribution π, observation emission probability B_m and state transition probability matrix A_m as follows:

$$\pi_k = P(s_1 = z_k), k = 1, 2, \dots M_h \tag{16}$$

$$B_m = [b_{k,t}] = P(y_t = j | s_t = z_k, x_t) \tag{17}$$

$$A_m = [a_{k_1,k_2}] = P(s_{t+1} = z_{k_2} | s_t = z_{k_1}, s_{t-1}, \dots, s_{t-\ell_d^*}), 1 \le k_1, k_2 \le M_h \tag{18}$$

where π_k denotes the probability distribution over the hidden states s_1 at the initial time step, representing the participant's initial cognitive condition. The emission probability function $b_{k,t}$ defines the probability of generating an observation y_t (the decision made at time step t, represented by the identifier of the next chosen node), given the hidden state z_k and acquired spatial information $x_t = [L_d, G_d, V_r, V_s, V_o, I_1, I_2, I_3, I_4, I_5, H_1, H_2, H_3, M_t, S_t \| E_{j'}]_t$. The spatial information x_t encompasses the participant's perceived distance and visual cues at time t, along with individual attributes, spatial encoding of the current node, travel purpose (as derived in Section 4.1), and interaction terms (identified in Section 4.2). In addition, $E_{j'}$ denotes the spatial feature encoding of the node j' visited at the previous time step. $[\cdot \| \cdot]$ represents the concatenation operation.

The model adopts a non-fixed mapping between hidden states and the Logit model specification, allowing the same latent state to be influenced by different factors across time steps. This design enables the capture of dynamic heterogeneity in decision-making behavior. Within each hidden state, the conditional probability of selecting a specific node–given the spatial information available at the current time step–is modeled using a mixed Logit formulation. Let $\beta_{k,t}$ denote the dynamic parameter vector associated with hidden state s_t at time t. Given the feature vector x_t and the current latent state $s_t = z_k$, the conditional probability of selecting decision node y_t at time t is expressed as:

$$P(y_t = j \mid x_t, s_t = z_k) = \frac{exp\left(\beta_{z_k,t,j}^T x_t\right)}{\sum_{(j')} exp\left(\beta_{z_k,t,j'}^T x_t\right)} \tag{19}$$

where $\mathcal{N}(j')$ denotes the set of adjacent nodes to node j in graph $G = (V, E)$, where j is the node currently occupied by the participant.

Modeling participants' dynamic wayfinding decision process using the DHMM-Logit framework can be decomposed into two subproblems: learning and decoding. The parameter estimation approaches for the DHMM-Logit model under these two settings are described as follows.

1) Learning problem. This problem is defined as maximizing the log-likelihood function $\mathcal{L}(\Theta)$ of the observed decision sequence $(y_1, y_2, \dots, y_T)$, given the corresponding feature sequence $(x_1, x_2, \dots, x_T)$ by estimating the model parameters $\Theta = \{\pi, A_m, \beta\}$.

$$L(\Theta) = \ln P(y \mid x, \Theta) = \ln \sum_s \left[\pi_{s_1} \prod_{t=\ell_d^*+1}^{T'} A_{m,s_t,s_{t+1}} \prod_{t=1}^{T'} P(y_t \mid x_t, s_t; \beta) \right] \tag{20}$$

The study employs the Expectation-Maximization (EM) algorithm to estimate the parameters of the DHMM-Logit model under the learning problem setting. In the Expectation step (E-step), the forward-backward algorithm is utilized to compute the forward probabilities, backward probabilities, and posterior probabilities of both states and state transitions based on the current model parameters.

This allows the estimation of the probability distribution over hidden cognitive states at each observation time step. In the subsequent Maximization step (M-step), the posterior probabilities obtained in the E-step are used to update the model parameters, including the initial hidden state distribution, the state transition probability matrix, and the parameters of the dynamic Logit models associated with each latent state. The update process is guided by the maximization of the log-likelihood function of the observed sequence. This iterative procedure alternates between the E-step and M-step until convergence is achieved, defined by a change in the log-likelihood smaller than a predefined threshold (set to 1×10^{-3}). Upon convergence, the resulting parameter estimates provide a stable characterization of the model. The estimated parameters obtained from the EM algorithm serve to reveal the temporal evolution of participants' latent cognitive states within the integrated transport hub, as well as the state-dependent influences of spatial and personal information on route choice decisions. This contributes to a deeper understanding of the heterogeneity in cognitive states and the dynamic mechanisms underlying pedestrian decision-making behavior over time.

2) Decoding problem. This problem is a continuation of the learning task and is fundamentally defined as a decoding problem. Given the estimated model parameters $\Theta = \{\pi, A_m, \beta\}$, the observed decision sequence $(y_1, y_2, \ldots, y_{T'})$ and corresponding feature sequence $(x_1, x_2, \ldots, x_{T'})$, the objective is to determine the optimal hidden state sequence $\hat{s} = \{s_1, s_2, \ldots, s_{T'}\}$ that best explains the observed dynamics. This is achieved by maximizing the joint probability $P(s, y|x, \Theta)$.

The decoding problem is addressed using the Viterbi algorithm, which infers the most likely hidden state sequence $\hat{s} = \{s_1, s_2, \ldots, s_{T'}\}$. Following the inference of the hidden state sequence, the study further performs prediction of participants' node selection behavior at each decision point. To enhance predictive robustness and mitigate the limitations associated with relying on a single inferred state, the study proposes a posterior-weighted ensemble prediction method based on the hidden state posterior probabilities. This method performs posterior-weighted fusion of the prediction probabilities across all hidden states at each time step, thereby mitigating the limitations associated with relying on a single inferred state.

$$\arg\max P(y_t = \hat{y}_t \mid x_t, \mu_t) = \sum_{k=1}^{M_h} \mu_t(z_k) \cdot P(y_t \mid s_t, x_t, \beta) \tag{21}$$

where, $\mu_t(z_k)$ denote the posterior probability of being in hidden state z_k at time step t, obtained via the forward-backward algorithm, and let $\hat{y}_t$ represent the predicted node number. The advantage of this strategy lies in its utilization of the full probabilistic distribution over hidden states rather than relying solely on the most probable single state (i.e., hard classification). By doing so, it avoids the potential loss of information caused by ignoring less likely but relevant cognitive states, thereby enhancing the robustness and accuracy of decision prediction in dynamic wayfinding scenarios. The complete decoding process is shown in Algorithm 1.

The final inferred hidden state sequence reveals the latent cognitive state transition patterns experienced by participants during the actual wayfinding process. This provides a foundational basis for analyzing and comparing the cognitive state dynamics of pedestrians arriving via different transportation modes within an integrated transport hub, as well as for quantifying the heterogeneity in decision-making behaviors across passenger groups.

where, $\mu_t(z_k)$ denote the posterior probability of being in hidden state z_k at time step t, obtained via the forward-backward algorithm, and let $\hat{y}_t$ represent the predicted node number. The advantage of this strategy lies in its utilization of the full probabilistic distribution over hidden states rather than relying solely on the most probable single state (i.e., hard classification). By doing so, it avoids the potential loss of information caused by ignoring less likely but relevant cognitive states, thereby enhancing the robustness and accuracy of decision prediction in dynamic wayfinding scenarios. The complete decoding process is shown in Algorithm 1.

The final inferred hidden state sequence reveals the latent cognitive state transition patterns experienced by participants during the actual wayfinding process. This provides a foundational basis for analyzing and comparing the cognitive state dynamics of pedestrians arriving via different transportation modes within an integrated transport hub, as well as for quantifying the heterogeneity in decision-making behaviors across passenger groups.

5. Modelling results and findings

This chapter presents the key research findings on pedestrian wayfinding behavior in integrated transport hubs based on the DHMM-Logit framework. Section 5.1 first reports the results of interaction term selection and significance analysis using tree-based models. Section 5.2 then introduces the computed outcomes of three indicators related to decision cascading in wayfinding behavior. Finally, Section 5.3 presents the calibrated results of the dynamic influence of explanatory variables on decision-making, along with an analysis of the evolution of hidden cognitive states throughout the wayfinding process.

5.1. Interaction term selection results

The RF model used the Gini index as the criterion for node splitting. Model hyperparameters were optimized through a combination of 5-fold cross-validation (K = 5) and grid search. The optimal number of trees was determined based on the 1-SE rule, as the out-of-bag (OOB) error learning curve showed that model accuracy plateaued when the number of trees exceeded 300. The hyperparameter configuration adopted in the study for implementing the interaction significance testing based on RF and SHAP is as follows:

- Number of permutations: B_r : 300
- Number of top features selected for main effect screening k : 5
- Number of trees in RF B : 100
- Significance level threshold: 0.05

Algorithm 1 First-order HMM-logit viterbi with posterior-fusion.

Require: M_h (number of states), transition matrix $A_m = \{a_{j,k}\}$, initial distribution $\Pi = \{\pi_k\}$, transition probabilities $A_m = \{a_{j,k}\}$
observations $O = \{o_t\}_{t=1}^{T}$, features $X = \{x_t\}_{t=1}^{T}$
Logit models $\{\text{model}_{k,b}\}$
Ensure: hidden path $S^* = \{s_t^*\}_{t=1}^{T}$, fused predictions $\hat{Y} = \{\hat{y}_t\}_{t=1}^{T}$
Compute emission log-likelihoods
for $t = 1$ **to** T **do**
  for $k = 1$ **to** M_h **do**
    $e[t,k] \leftarrow \log(\text{model}_{k,b}.\text{predict_proba}(x_t)[o_t])$
  end for
end for
Initialize
for $k = 1$ **to** M_h **do**
  $\delta[1,k] \leftarrow \log(\pi_k) + e[1,k]$
  $\psi[1,k] \leftarrow 0$
end for
Viterbi recursion
for $t = 2$ **to** T **do**
  for $k = 1$ **to** M_h **do**
    $(\delta[t,k], \psi[t,k]) \leftarrow \max_{j=1\ldots M_h}\{\delta[t-1,j] + \log(a_{j,k})\}$
    $\delta[t,k] \leftarrow \delta[t,k] + e[t,k]$
  end for
end for
Termination
$s_T^* \leftarrow \arg\max_{k=1\ldots M_h} \delta[T,k]$
Back-trace
for $t = T-1$ **downto** 1 **do**
  $s_t^* \leftarrow \psi[t+1, s_{t+1}^*]$
end for
Compute posteriors $\mu_t(k) = P(s_t = k \mid O, X)$ via forward-backward
Posterior-fusion prediction
for $t = 1$ **to** T **do**
  for j **in** all candidate labels **do**
    $p_t(j) \leftarrow \sum_{k=1}^{M_h} \mu_t(k)\text{model}_{k,b(t)}.\text{predict_proba}(x_t)[j]$
  end for
  $\hat{y}_t \leftarrow \arg\max_j p_t(j)$
end for
return $S^*, \hat{Y}$

Table 3
Statistics on interaction items with significant interaction intensity. The study used different numbers of asterisks (*) to indicate differences in significance levels.

Interaction item	Interaction intensity	p value	AIC	BIC
$I_4 \bullet H_1$	2.59	0.0037**	9515.56	15174.14
$L_d \bullet H_1$	0.01	0.0158*	9536.76	15195.35
$I_1 \bullet I_4$	0.01	0.0032**	9565.76	15224.35
$V_o \bullet I_4$	−1.27	<0.0001***	9663.48	15322.06
$I_3 \bullet H_1$	−0.02	<0.0001***	9753.91	15412.5
$I_4 \bullet H_3$	−2.81	0.0001***	9823.16	15481.74
$I_3 \bullet H_3$	0.01	<0.0001***	9824.27	15482.86

The seven interaction terms that exhibit both statistical significance (p-value < 0.05) and strong interaction intensity (interaction intensity 0.01) with respect to decision-making behavior are summarized in Table 3.

From the perspective of interaction semantics, a strong positive interaction (2.59) is observed between Visual Clustering Coefficient and the use of 2D schematic navigation, indicating that participants using 2D maps in hub environments tend to leverage global spatial information and preferentially choose paths with higher visual clustering. Such paths typically exhibit clearer structural coherence and are easier to navigate. This synergy suggests that the use of 2D maps enhances the effectiveness of visually cohesive regions in guiding wayfinding decisions. In contrast, Local Distance Information shows a mild positive interaction (0.01) with 2D map usage. This implies that participants using 2D maps may adjust their sensitivity to local distances during the wayfinding process, relying

less on direct spatial proximity between nodes compared to those navigating without maps. Conversely, Isovist Perimeter exhibits a negative interaction coefficient (-0.02) with 2D map usage. This suggests that the presence of additional spatial information from the schematic map reduces participants' reliance on immediate visual cues, thereby significantly altering their navigational strategy.

For interaction terms related to early arrival duration, a strong negative interaction (-2.81) is observed with the Visual Clustering Coefficient. This suggests that participants with longer early arrival times are less inclined to prioritize paths with high visual clustering, potentially due to lower time pressure, which permits more exploratory behavior during the wayfinding process. In contrast, Isovist Perimeter exhibits a slight positive interaction (0.01) with early arrival duration, indicating that participants who arrive earlier tend to be more sensitive to paths with higher visual complexity. This finding aligns with the pattern observed in interaction term $I_4 \bullet H_3$, which suggests that passengers with extended early arrival times demonstrate more exploratory spatial behavior. In conjunction with the spatial syntax visualization of the hub, this result supports the interpretation that early-arriving passengers are more likely to engage in exploratory activities along the peripheral corridors of the hub, areas characterized by high Isovist Perimeter and low Visual Clustering Coefficient. Conversely, passengers with limited early arrival durations tend to take more direct routes to their destination by navigating through key junctions that lead to the waiting area, such as nodes 11 and 13.

To further quantify the impact of incorporating interaction terms on model predictive performance, the study employs Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) as evaluation metrics. The results indicate that the model including the interaction between Visual Clustering Coefficient and 2D schematic navigation usage achieves the lowest AIC (9515.56) and BIC (15174.14), suggesting this interaction contributes most significantly to improving predictive accuracy. The second-best performance is observed for the interaction between Local Distance and 2D schematic navigation usage, with AIC = 9536.76 and BIC = 15195.35, indicating a moderate yet meaningful enhancement in model fit. In contrast, although the interaction between Isovist Perimeter and early arrival duration is statistically significant, it yields the highest AIC (9824.27) and BIC (15482.86), suggesting its marginal contribution to overall model predictive performance is relatively limited. Based on the AIC rankings, the top three interaction terms with the lowest AIC values are selected and incorporated into the explanatory variable set of the subsequent DHMM-Logit modeling stage.

5.2. Decision cascade effect analysis

This section aims to conduct a quantitative analysis of the cascading effects embedded in pedestrian wayfinding decision sequences, based on the theoretical framework established in Section 4.3 on behavioral decision cascading modeling. Specifically, by utilizing the Mutual Information (MI) distribution heatmap (see Fig. 11, left), the information decay curve (Fig. 11, right), the decision entropy distribution (see Fig. 12), and the attribute-based heterogeneity analysis (see Table 3), we assess the temporal dependencies and the dynamic evolution of uncertainty across decision steps. This analysis serves not only to empirically validate the temporal cascading hypothesis underpinning hidden state transitions in the DHMM-Logit model, but also to reveal key patterns such as the exponential decay of decision dependency, task-driven heterogeneity, and the modulatory effects of passenger attributes on cascading strength. By leveraging these indicators, we quantitatively determine the maximum influence duration of cascading effects and the appropriate number of hidden states, thereby providing a principled basis for model parameter optimization.

We begin by analyzing the mutual information (MI) decay in the decision sequence. As illustrated in Fig. 11, the left panel presents a pairwise MI matrix heatmap, where rows and columns represent subsequent and preceding decision steps, respectively. The color intensity indicates the MI strength, normalized to the range $[0, 1]$. The heatmap, aggregated over the full dataset, shows that MI values are highest near the main diagonal, indicating the strongest dependency between adjacent decision steps, $\ell_d = 1$. As the decision lag increases, the MI values exhibit an exponential decay pattern, as depicted in Fig. 11 (right). The MI drops below the predefined threshold of 0.1 after the 7th lagged step, leading to the estimation of the maximum influence duration $\ell_m^* = 6$. Furthermore, the decay curve reveals a half-life of approximately 2.5 decision steps, yielding a maximum significant influence length $\left\lfloor \ell_d^* \right\rfloor = 2$. Based on these findings, the order of the HMM is preliminarily constrained to the range of first-order to second-order. The final selection of the optimal model order will be further evaluated in Section 5.3, taking into account both model fit and predictive performance.

The heatmap and fitted decay curves jointly reveal significant differences in the mutual information decay patterns of pedestrian decision sequences under different arrival modes. While the overall decay trends are similar across modes, the MI matrix associated with metro arrivals exhibits notably stronger long-range dependencies. Specifically, regions beyond the main diagonal show darker color intensities, indicating that although the initial mutual information strength is comparable across all modes, the retained MI at lag 12 for metro arrivals remains around 4%, whereas it falls below 2% for road and air arrivals. Moreover, the decay rate for metro-arrival sequences is significantly slower. The exponential decay rate fitted to metro data is 2.2% and 12.8% lower than that for road and air arrivals, respectively. This suggests a more persistent influence of early decisions on subsequent wayfinding behavior among subway users. It is necessary to provide comprehensive and detailed wayfinding information at the transfer entrance of the hub (the first decision point) for arriving subway passengers. This should include floor plans, directions to the nearest turnstiles/ticket machines, and other relevant details to facilitate optimal travel decisions at the initial location.

We attribute this slower decay in the metro case to two main factors. First, the proximity of metro entry points to core hub facilities, reducing spatial uncertainty early in the journey. Second, the higher density and clarity of signage in metro-access areas, which reinforces initial directional decisions and sustains their influence over longer time spans. In contrast, the faster decay observed for road arrivals may result from the spatial dispersion of road entry points and the visual obstructions encountered upon entry. Visual analysis indicates that the obstacle coverage in the visual field at road entrances is 25% higher than that of metro and air entrances. Additionally, from a spatial syntax perspective, these areas exhibit lower Connectivity values, offering fewer intuitive decision cues

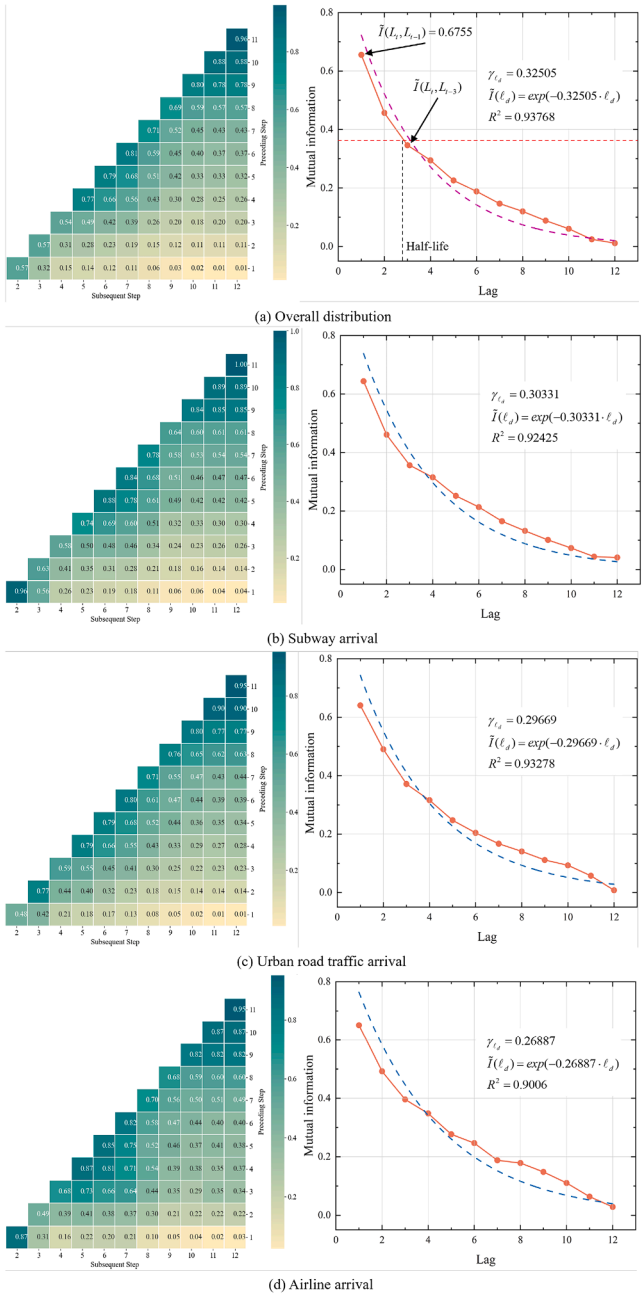


Fig. 11. Decision cascade mutual information (MI) analysis. Left: Pairwise MI matrix across steps. Shows the normalized mutual information strength from each preceding step (y-axis) to each subsequent step (x-axis). Right: MI decay curve with exponential fit. The solid red line depicts the overall mutual information decay as a function of lag (Lag), while the dashed blue line shows the exponential model fit, annotated with the fitted decay parameter γ_{ℓ_d} and the coefficient of determination R^2.

and lacking coherent guidance systems. As a result, travelers arriving by road tend to rely more on immediate environmental cues, leading to more reactive and less structured decision-making behavior.

As an important dimension of participant heterogeneity in this study, the use of 2D schematic navigation also warrants quantitative analysis regarding its effect on the decay of historical decision information. As shown in Fig. 12, the use of 2D navigation significantly mitigates the rate of mutual information decay during the wayfinding process, particularly in the early stages. Specifically, in the absence of 2D navigation, the mutual information retained at the third decision step declines to just 11.6% of its initial value. In

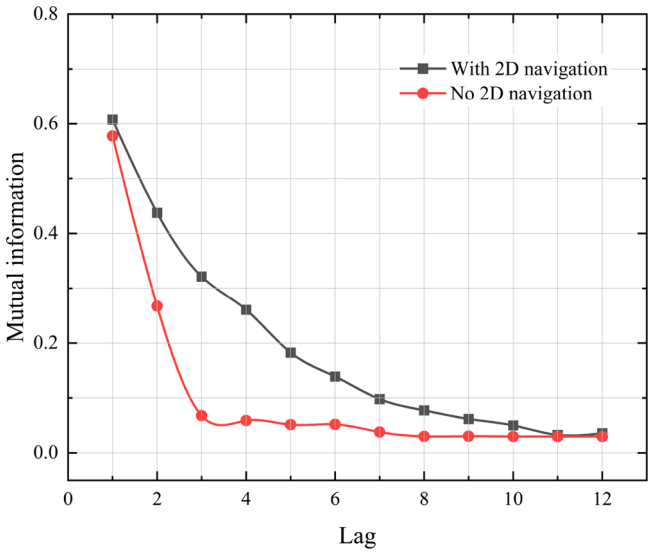


Fig. 12. Comparison of decision mutual information decay differences between navigating and non-navigating populations.

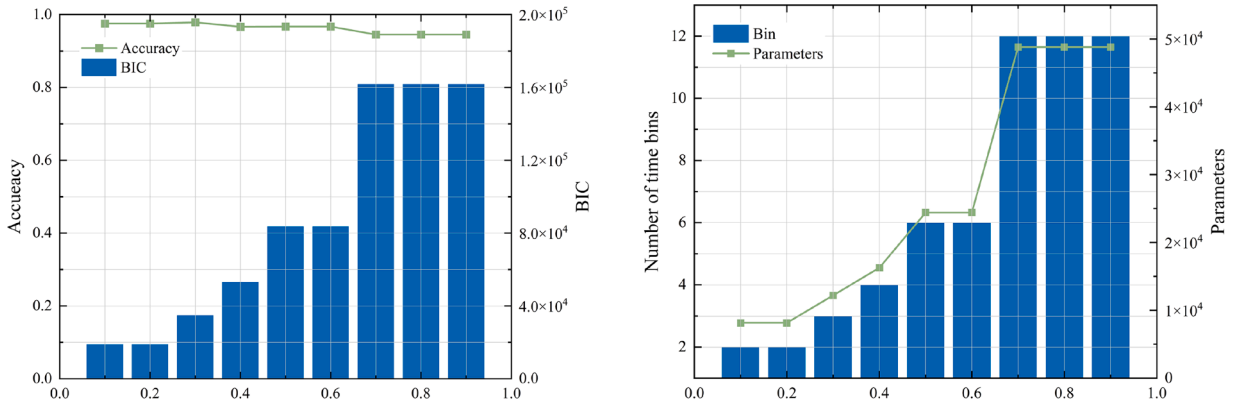


Fig. 13. Comparison of decision mutual information decay differences between navigating and non-navigating populations.

contrast, participants utilizing 2D navigation experience a substantially slower loss of decision-related information, reducing the decay gap by 41.2% at the same step.

These findings highlight the stabilizing effect of navigational aids on cognitive continuity, suggesting that 2D navigation enhances the retention of prior decision logic and contributes to more consistent and structured decision-making sequences.

Fig. 13 presents a comparison of the number of temporal bins, DHMM-Logit parameter complexity, and model fitting performance under varying threshold values τ_d. As the threshold increases, the model complexity reflected in the number of estimated parameters rises sharply. However, this increase does not correspond to a significant improvement in predictive accuracy. Notably, when $\tau_d \geq 0.4$, the prediction accuracy of the DHMM-Logit model begins to decline, and the Bayesian Information Criterion (BIC) increases rapidly, indicating overfitting and diminished generalization ability. Based on these observations, the threshold is set to 0.2, resulting in two temporal bins for the discretization of the decision sequences. This setting balance model parsimony and predictive performance, enabling effective capture of temporal dynamics while maintaining robustness.

The analysis of decision entropy distributions across different travel tasks reveals clear differentiation in behavioral uncertainty. Overall, the entropy values vary significantly by task type. The lowest entropy is observed in the post-shopping return to the waiting area task $H_e = 0.15$. This low entropy reflects the high determinism of the task, as it typically involves a simple reversal of the previous path, requiring minimal decision-making effort. The next lowest entropy values correspond to the waiting and ticket collection tasks. These are largely governed by standardized procedures (e.g., fixed locations of ticket vending machines), resulting in a narrow decision space. Furthermore, the waiting areas are typically characterized by high visual connectivity and spatial recognizability, which reduce uncertainty during navigation. In contrast, the highest entropy is associated with the shopping task $H_e = 1.72$, indicating significant behavioral heterogeneity. This arises from the dispersed spatial distribution of retail locations and the influence

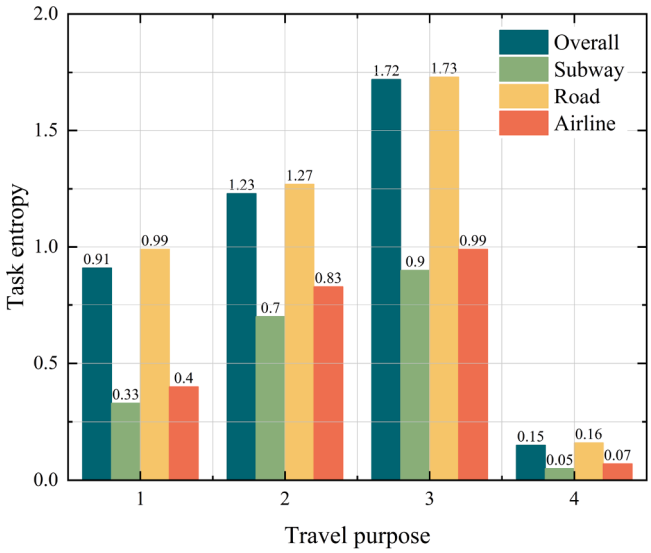


Fig. 14. The bar chart shows the distribution of task entropy among pedestrians arriving at the hub via different modes of transport.

Table 4
The impact of heterogeneity attributes on decision entropy in different wayfinding tasks.

Heterogeneity attribute	Value	Entropy for purposes				Sig.
		1	2	3	4	
H_1	0	0.94	0.99	1.33	0.11	0.001***
	1	0.69	1.42	1.73	0.07	
H_2	0	0.89	1.32	1.72	0.10	0.941
	1	0.93	0.98	1.57	0.08	

of individual preferences, which increase variability and reduce predictability in route choice. The overall task entropy computed across all sequences is 0.73. Based on this entropy level, the number of hidden states M_h in the DHMM-Logit model is set to 3.

Further comparison of task entropy distributions across different arrival modes reveals notable differences, as shown in Fig. 14. Pedestrians arriving via metro consistently exhibit lower decision entropy across all tasks (green bars), compared to those arriving by airline or road. Quantitatively, the task entropies for metro arrivals are 17.5%/66.7%, 15.7%/44.9%, 9.0%/47.9%, 28.6%/68.8% lower than air and road arrivals. In contrast, road arrivals (yellow bars) exhibit the highest entropy levels across all tasks, indicating greater behavioral uncertainty. This heightened entropy may stem from the spatial dispersion of road entrances, visual obstructions, and weaker guidance infrastructure, which together reduce the standardization of pedestrian behavior and lead to more varied, less predictable navigation strategies. These findings emphasize the modulatory role of arrival mode on task-related decision complexity, reinforcing the need to account for access-mode heterogeneity in cognitive state modeling.

Finally, the analysis is extended by examining task entropy differences across heterogeneous passenger attributes, as detailed in Table 4. The table reports the entropy values for different wayfinding tasks under two key heterogeneity dimensions: use of 2D navigation and spatial familiarity, along with the statistical significance of attribute-related effects on behavioral uncertainty. Quantitative results show that participants not using 2D navigation exhibit an average task entropy of 0.843, whereas those using 2D navigation demonstrate a 16.0% higher average entropy of 0.978. This suggests that while 2D navigation tools help users develop a more comprehensive understanding of spatial layout, they also increase the number of perceived alternative paths, thus expanding the decision space and elevating overall task entropy. For future travelers utilizing two-dimensional dynamic navigation, it is essential to not only provide electronic navigation services but also explicitly highlight clearly discernible shortest paths. This approach reduces travel decision-making complexity, thereby achieving the goal of guidance or route recommendation.

More specifically, for waiting and returning from shopping, 2D navigation reduces task entropy by 26.6% and 36.4%, respectively. In contrast, it increases entropy for ticket collection and shopping tasks by 43.4% and 30.1%, respectively. This dual effect likely stems from the nature of the tasks. For goal-directed, necessity-based tasks (e.g., waiting), where the destination is fixed and well-known, 2D navigation reduces cognitive load and narrows route choices, thereby decreasing decision uncertainty. Conversely, for personalized or exploratory tasks (e.g., shopping and self-service ticketing), where service facilities are more spatially dispersed, the navigation tool reveals a wider range of alternatives, increasing branching possibilities and inducing greater choice variability–resulting in a more uniform probability distribution and thus higher entropy.

Table 5
Comparison of DHMM-Logit and baseline model goodness-of-fit metrics.

Model	Accuracy	Log-likelihood	BIC
MXL	89.31%	−793.98	24057.32
Second order HMM-Logit (3 hidden state)	91.98%	−659.52	23113.60
First order DHMM-Logit (2 hidden state)	97.03%	−640.29	19470.12
DHMM-Logit (4 hidden state)	95.41%	−605.16	19960.20
DHMM-Logit (No global distance)	95.28%	−651.50	19761.53
DHMM-Logit (One-hot spatial encoding)	94.24%	−659.91	19609.35
DHMM-Logit (Original Viterbi)	92.35%	−701.47	20134.21
DHMM-Logit (Visual depth → L_t)	91.07%	−791.56	21576.79
First order DHMM-Logit (3 hidden state)	**97.54%**	**−585.34**	**18850.64**

These findings suggest that while 2D navigation systems can enhance efficiency in hub navigation, they may also amplify decision uncertainty in highly individualized wayfinding scenarios. Therefore, it is recommended that task-specific optimization be considered in navigation interface design, for instance, simplifying or constraining the decision space for shopping-related tasks to reduce cognitive burden and improve user experience.

To verify the statistical significance of heterogeneous attributes on participants' decision-making behavior, the study employs a sequence-level similarity analysis. Specifically, decision sequences from participants with different attribute values (e.g., 2D navigation usage) are concatenated into vectors and compared using the Levenshtein distance (Lcvenshtcin, 1966), which quantifies the minimum number of edits required to transform one sequence into another. The test results show that 2D navigation usage has a statistically significant effect on decision-making behavior, with a p-value of 0.001 (<0.05).

Differences in spatial familiarity similarly leads to variations in task entropy. For participants with no prior familiarity with the environment (familiarity = 0), the average task entropy is 1.008, which is 11.7% higher than that of participants with full familiarity (familiarity = 1), whose average entropy is 0.89. This effect is particularly notable in the ticket collection and shopping tasks, where familiar participants exhibit entropy reductions of 0.34 and 0.15, respectively. These results suggest that familiarity with the environment enables participants to leverage prior knowledge (e.g., mental maps or memory of layout) to reduce exploratory demands, thereby lowering decision uncertainty in most tasks. However, the PERMANOVA test indicates that the overall effect of spatial familiarity on behavioral decisions is not statistically significant (p-value$=0.941>0.05$). This implies that, within the current dataset, the influence of spatial familiarity on decision-making is relatively weak and fails to produce a detectable divergence in sequence-level patterns.

5.3. Analysis of prediction performance and parameter estimation results of DHMM-Logit

5.3.1. Model prediction performance analysis and ablation experiments

During model training, the dataset is split into a training set (70%) and a test set (30%). To evaluate the performance of our dynamic discrete choice model, we compare it against several benchmark models, including the Mixed Logit (MXL) model, first-order HMM-Logit models with 2 and 3 hidden states, and a second-order HMM-Logit model with 3 hidden states. These models are assessed based on goodness-of-fit metrics and through ablation studies to validate the contribution of key components to predictive performance. The comparison focuses on three core evaluation indicators, including prediction accuracy, log-likelihood and BIC. To quantify prediction performance, we define prediction accuracy of pedestrian wayfinding decisions as:

$$Accuracy = \frac{1}{\sum_{i=1}^{N} T'_i} \sum_{i=1}^{N} \sum_{t=1}^{T'} f_z(\hat{y}_{i,t} = y_{i,t}) \tag{22}$$

where N is the total number of decision sequence. T'_i represents the length of i^{th} sequence. $f_z(\cdot)$ denotes the indicator function, when the predicted result matches the actual result, the value is 1. Otherwise, the value is 0.

Table 5 presents a comparative analysis of the model fit indices between the proposed DHMM-Logit model and baseline models. The results indicate that the first-order DHMM-Logit consistently outperforms the benchmark models across all evaluation metrics. Specifically, the first-order DHMM-Logit achieves substantially higher log-likelihood, lower Bayesian Information Criterion (BIC), and greater predictive accuracy than both the Mixed Logit (MXL) model and the second-order HMM-Logit model. Furthermore, in terms of BIC and log-likelihood, the DHMM-Logit model with three latent states exhibits a lower BIC and a higher log-likelihood compared to its two-state counterpart, suggesting that the three-state specification provides the best fit to the observed data.

The ablation study results demonstrate that the improved Viterbi decoding algorithm enhances node prediction accuracy by 5.19%. Although incorporating node spatial feature embeddings derived from the Skip-gram model as explanatory variables in the DHMM-Logit framework yields a relatively modest improvement in accuracy (3.3%), it significantly reduces the Bayesian Information Criterion (BIC) by 31.72% and increases the log-likelihood by 11.30%. The research findings also demonstrate that considering global distance is effective in enhancing the prediction of passenger wayfinding behavior (accuracy increased from 95.28% to 97.54%.). Moreover, aligning the visual depth information with the subsequent target node L_{t+1} (represented by Visual depth → L_{t+1}) leads to a 6.47% increase in prediction accuracy. These findings collectively validate the effectiveness of the proposed model enhancements and data preprocessing strategies in improving the performance of the DHMM-Logit framework.

Table 6
K-fold cross-validation grouped by participants.The evaluation metrics presented in the table include: Accuracy (calculated by weighting and aggregating the posterior model parameters obtained through the forward-backward algorithm for each state and influencing factor to derive the predicted category at each step, which is then compared with the true next node), Log-likelihood (computed on the test set using the trained parameters and emission probability matrix through scaled forward-backward recursion), and BIC (where a smaller value indicates that higher out-of-sample goodness of fit is achieved with lower model complexity in that fold).

Fold	Accuracy	Log-likelihood	BIC	Number of participants
1	95.48%	−372.55	18318.90	10
2	92.13%	−452.76	16923.37	10
3	92.62%	−471.32	17954.34	10
4	96.52%	−357.22	18297.53	10
5	95.36%	−468.66	15607.76	10
Average	94.42%	−424.50	17420.38	50
SD	1.93%	49.33	1038.72	–

Table 7
Initial hidden state probability distribution matrix for groups of individuals with different heterogeneous attributes. The study verified the significance of the impact by conducting a chi-square test on the frequency distribution of hidden states under different values of heterogeneous attributes.

Heterogeneity attribute	Value	Hidden states 1	Hidden states 2	Hidden states 3	**sig.** (χ^2)
H_1	0	0.52	0.37	0.11	0.577
	1	0.51	0.40	0.09	
H_2	0	0.48	0.42	0.10	0.751
	1	0.53	0.37	0.10	
H_3	1	0.44	0.47	0.09	0.809
	2	0.50	0.39	0.11	
	3	0.54	0.38	0.08	
	4	0.58	0.25	0.17	
Transfer mode	Subway	0.56	0.29	0.15	$1.979 \times 10^{-18**}$
	Road	0.42	0.46	0.12	
	Airline	0.20	0.78	0.02	

Since multiple activity chains correspond to the same participant and exhibit clustered dependence, constructing the training dataset solely based on temporal segmentation leads to information leakage within travel chains and hinders the model's generalization ability for predicting overall participant behavior. Therefore, we replaced the original 7:3 sequence-level split with participant-grouped K-fold cross-validation, ensuring that all activity chain data corresponding to the same participant are consistently assigned to the same fold. The model's predictive accuracy, log-likelihood, and BIC were evaluated on the training folds, and the results are reported in Table 6. The prediction results indicate that under participant-grouped conditions, the average prediction accuracy remains above 94.4%. The standard deviations of the three-evaluation metrics across folds (Accuracy: 1.93%, Log-likelihood: 49.33, BIC: 1038.72) demonstrate that the model maintains robustness in out-of-sample evaluation and achieves a good overall fit.

5.3.2. Regression results and hidden state evolution analysis

This section first presents the initial probability distribution of latent states in the pedestrian wayfinding process, considering the influence of individual heterogeneity and variations in arrival modes. It then reports the regression coefficients of explanatory variables estimated under each latent state using the dynamic Logit model. From an evolutionary perspective, the analysis explores the relationship between the temporal transitions of latent state sequences and the dynamics of variable effects. By examining the differences in estimated coefficients across the three latent states, the study elucidates how pedestrians with distinct heterogeneous attributes exhibit divergent spatial cognition patterns during wayfinding in the hub environment.

Table 7 presents the initial probability distribution across three latent states for participant groups with varying heterogeneity attributes. The mode of arrival within the hub exerts a significant influence on the initial state distribution (Chi-square test, $\chi^2 = 41.642$, $p = 1.979 \times 10^{-18} < 0.05$), highlighting the mode-specific wayfinding strategies that emerge due to the spatial heterogeneity of multimodal transfer layouts. Nearly half of the metro-arriving pedestrians are initially assigned to State 2, a tendency that is even more pronounced among air travelers (78%), whereas participants-arriving by road exhibit a more balanced distribution between States 1 and 2. In contrast, the use of two-dimensional (2D) navigation does not significantly affect the initial latent state distribution (Chi-square test, $\chi^2 = 1.099$, $p = 0.577$). Similarly, neither spatial familiarity ($\chi^2 = 0.575$, $p = 0.751$) nor early arrival time ($\chi^2 = 0.821$, $p = 0.809$) shows a significant effect. These findings suggest that the initial latent state is primarily shaped by the objective spatial configuration and access point locations within the hub, while subjective factors such as digital navigation aids and prior familiarity with the environment exert limited influence on passengers' initial spatial cognition.

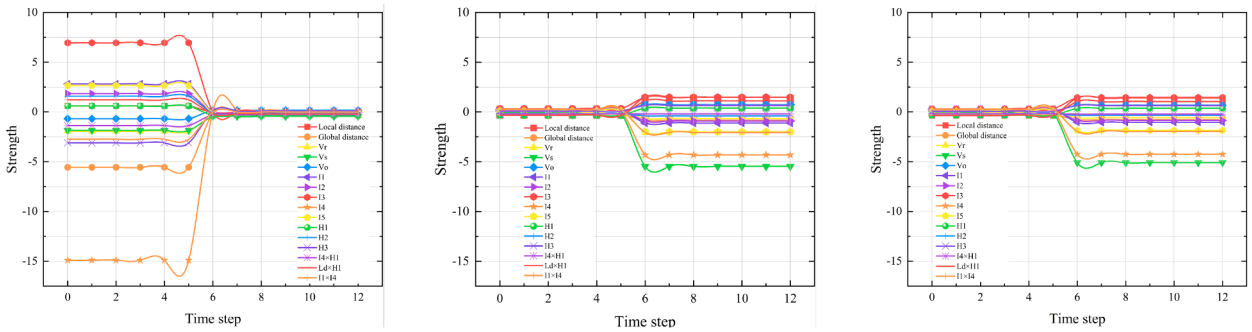


Fig. 15. Line chart of dynamic regression coefficients for overall participant groups in different hidden states (state 1 (a), state 2(b), state 3(c)).

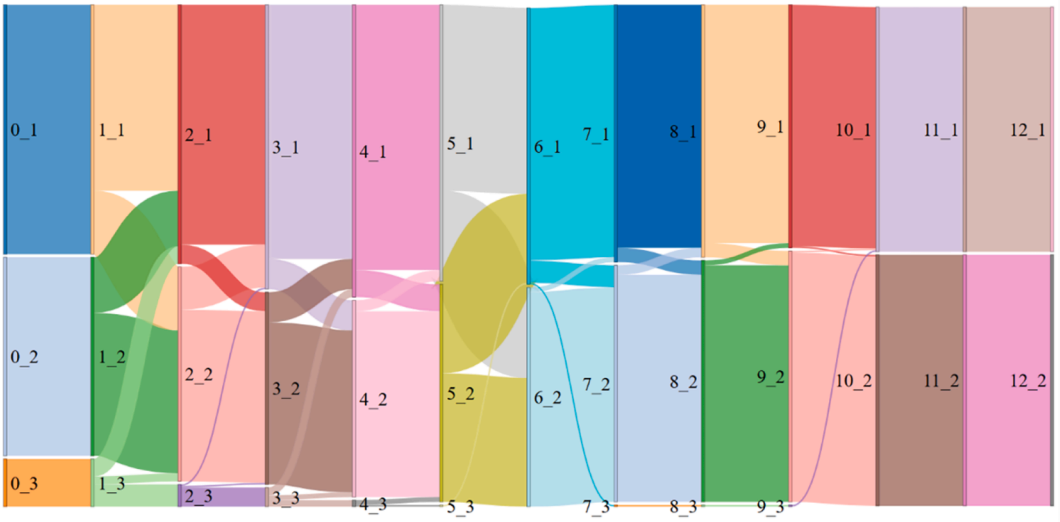


Fig. 16. Sankey diagram comparing hidden state transitions of overall participant groups.

The regression results of the DHMM-Logit model presented in Figs. 15 and 16 reveal pronounced distinctions among the three latent states. By jointly analyzing the state transition frequencies and initial state distributions, the behavioral interpretations of each latent state can be inferred. Notably, States 2 and 3 exhibit highly similar coefficient magnitudes and trends during the second time interval, suggesting overlapping decision-making patterns at this stage. In contrast, State 1 demonstrates significantly divergent coefficient dynamics, indicating a distinct behavioral profile. Transition frequency analysis further shows that States 1 and 2 dominate the overall wayfinding process, accounting for approximately 50% and 40% of initial state occurrences, respectively. Beyond time step $t > 5$, both states exhibit strong self-transition probabilities exceeding 94%, indicating stabilized decision strategies. In contrast, State 3 appears predominantly in the early stages as a transient state with low frequency. It gradually fades out as the decision process evolves, transitioning into either State 1 or State 2. This pattern suggests that State 3 represents an unstable or exploratory phase in early wayfinding, while States 1 and 2 reflect more established navigation strategies.

The dynamic characteristics of regression coefficients across latent states reveal distinct behavioral patterns in pedestrian wayfinding. During the early phase (first temporal bin, $t = 0 - 5$), individuals in State 1 exhibit pronounced sensitivities to spatial and visual features, as evidenced by the relatively large positive coefficients for connectivity (2.81), isovist area (1.85), isovist perimeter (6.95), visual integration (2.66), familiarity (1.58), and 2D navigation usage (0.60). Conversely, negative coefficients are observed for global distance (−5.56), route/space visual occupancy (−1.94 / −1.85), obstacle visual occupancy (−0.68), visual clustering coefficient (−14.90), and ETA (−3.10). These results indicate that pedestrians in State 1 are highly responsive to locally available spatial-visual cues and prefer paths characterized by dense physical and perceptual information. Their navigation strategy can be interpreted as vision-guided exploration, favoring routes with optimal network connectivity and expansive isovist fields to avoid detours or time-related stress caused by unfamiliar layouts.

In contrast, States 2 and 3 display substantially weaker coefficient magnitudes. Despite retaining positive associations with obstacle visual occupancy (0.07/0.05), connectivity (0.24/0.25), isovist-based metrics (0.24-0.32 / 0.22-0.31), and ETA (0.09/0.08), these effects are marginal. Notably, distance-related variables, including local distance (−0.34 / −0.26) and global distance (−0.37 / −0.29), alongside 2D navigation usage (−0.33 / −0.33) and familiarity (−0.17 / −0.18), exhibit consistently negative coefficients.

Table 8
Regression coefficient fitting results for each variable under different travel purposes. Waiting behavior shows a preference for short travel distances and exhibits sensitivity to spatial information density along the path. Shopping behavior tends to favor more enclosed areas with smaller isovist fields. When returning to the waiting area after shopping, passengers strongly rely on clear and transparent visual pathways and prefer the shortest overall route compared to other activities (since this group of people has already travelled from the waiting area to the shopping area, their spatial cognition requirements are lower, and their overall planning behaviour patterns become apparent). Ticket collection is characterized by a tendency for proximity-based choices and displays insensitivity to spatial-geometric attributes. While 2D navigation significantly enhances the efficiency of shopping and return-to-waiting tasks, it hinders performance in waiting and ticket collection activities–likely because these simpler tasks suffer from cognitive overload when unnecessary information is introduced. Familiarity exerts a positive effect only on waiting behavior.

Misson	Proceed to the waiting area	Ticket collection	Shopping	Return to the waiting area
Intercept	3.836	−4.098	−1.808	2.069
L_d	−0.591	−0.623	0.368	−0.394
G_d	0.005	−0.038	0.337	−0.304
V_v	0.213	−0.284	0.161	0.409
V_s	0.023	−0.151	0.134	0.394
V_o	0.276	0.043	0.143	−0.763
I_1	−0.753	−0.126	−1.793	−0.713
I_2	−0.597	−0.169	−1.897	−0.730
I_3	0.864	−0.023	−0.012	0.171
I_4	0.747	0.042	−1.385	0.596
I_5	0.371	−0.107	−0.079	0.216
H_1	−0.950	−0.565	1.016	0.499
H_2	1.233	−0.047	−0.095	−0.035
H_3	−0.573	0.037	0.047	0.088
$I_4 \bullet H_1$	0.480	0.829	0.361	0.948
$L_d \bullet H_1$	0.222	0.296	0.049	0.123
$I_1 \bullet I_4$	0.271	0.069	0.484	0.143

This suggests a higher tolerance for local obstacles and a diminished reliance on visual-spatial cues, with a corresponding decline of 30%-50% in the explanatory strength of visual factors compared to State 1. These traits reflect a behavioral mode associated with low spatial awareness and limited efficiency orientation, likely stemming from the engagement in non-essential activities such as ticket collection and shopping. Statistical analysis confirms that 46% of individuals exhibiting such behaviors during the early stages are categorized in State 2. Their actions suggest a greater acceptance of spatial uncertainty and reduced urgency concerning travel time or route optimization.

Further evidence is provided in Table 8, which presents overall coefficient estimates stratified by travel purpose. It becomes evident that only for essential activities like waiting, visual factors are positively associated and exhibit high coefficient strengths. In contrast, for ticketing and shopping, visual predictors often carry negative coefficients of lower magnitude. Notably, ticketing behavior is primarily influenced by the shortest local distance, whereas shopping behavior shows similar sensitivity to both local and global distances, indicating broader variability in spatial preferences. These findings suggest that passengers engaged in non-essential, heterogeneous activities may willingly disregard efficiency-based criteria, including spatial proximity and visual salience, in favor of accessing personally meaningful destinations. This behavioral divergence results in attenuated or inverse effects of visual and distance-related features in their decision models.

In the later stage of the wayfinding process (second temporal bin, $t = 6 - 11$), the strategy associated with State 1 exhibits a clear trend toward simplification. Most visual variables, previously exerting strong positive effects, now display negative coefficients–for example, route visual occupancy (−0.33) and connectivity (-0.04). The coefficient for obstacle visual occupancy reverses from negative to positive (0.18), while visual integration–which had been a key positive predictor in early decision-making declines significantly from 2.66 to −1.99. This shift reflects an adaptation to local environmental constraints, as pedestrians gradually accumulate spatial knowledge and reduce their dependence on global visual cues. The observed transition from cautious exploration to goal-oriented efficiency illustrates the behavioral transformation driven by increased environmental familiarity, minimizing inefficient trial-and-error routes. This behavioral evolution is further supported by the state transition diagram for 2D navigation users (Fig. 17) and the mutual information decay analysis. These findings show that State 1 dominates among navigation-assisted and metro-arriving pedestrians, suggesting that the simplification of strategies in the later phase is a combined effect of digital navigation tool use and accumulated spatial familiarity under decaying information dynamics.

In contrast, States 2 and 3 converge in their decision patterns, with positive coefficients for Local Distance (1.11 / 1.05), Isovist Perimeter (1.47 / 1.45), and Obstacle Visual Occupancy (0.74 / 0.68). Simultaneously, negative coefficients are observed for Global Distance (−0.81 / −0.76), Visual Field Area (−5.46 / −5.09), Visual Clustering Coefficient (−4.31 / −4.25), Visual Integration (−1.99 / −1.85), and Familiarity (−0.39 / −0.34). Compared to the earlier stage, the sensitivity to local barriers and spatial constraints is significantly enhanced–with positive coefficients nearly doubling–while avoidance of global uncertainty and open spaces becomes

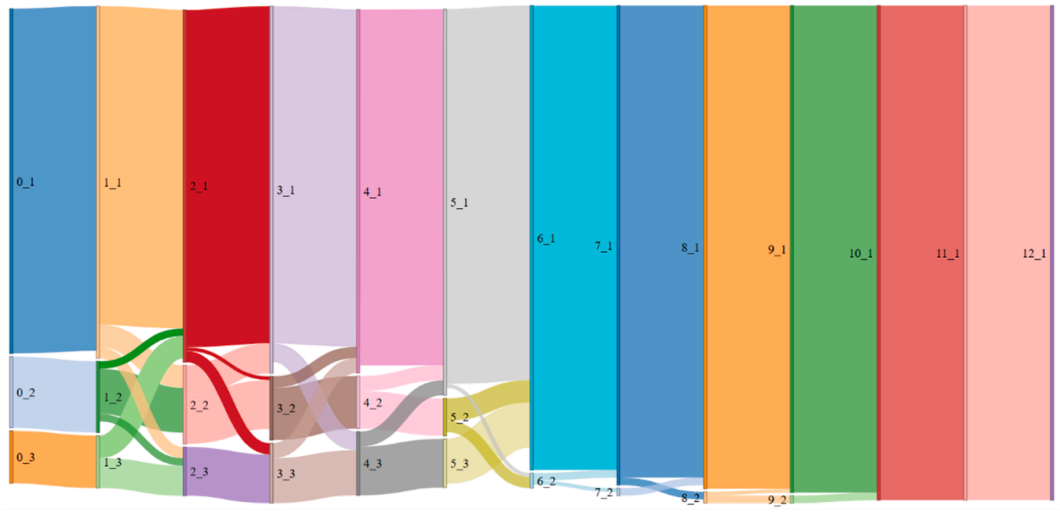


Fig. 17. Sankey diagram comparing hidden state transitions of participant groups using 2D navigation.

more pronounced, as evidenced by the highly negative coefficients for visual spatial area (e.g., −5.46 in State 2 and −0.33 in State 1). Moreover, visual cue influence in States 2 and 3 increases by 2-3 times relative to State 1, signifying a shift from visual insensitivity to high visual sensitivity. This pattern reflects a heuristic convergence strategy, where pedestrians particularly those not using navigation tools leverage fragmented edge structures and complex peripheral visual cues to minimize cognitive load while accurately identifying goal nodes. In doing so, they avoid reliance on full spatial representations and instead adopt efficient, visually grounded shortcuts within complex environments.

In summary, this phase marks a divergence in strategy: while navigation-assisted pedestrians in State 1 streamline decisions due to technological support and accumulated knowledge, non-assisted pedestrians in States 2 and 3 refine their reliance on environmental visual fragments to efficiently complete their tasks with minimal cognitive overhead.

The statistical analysis of state transition patterns further reinforces the understanding of the evolutionary dynamics underlying pedestrian wayfinding behavior. During the early stage, high transition probabilities are observed from State 3 to the other two states specifically, with State 3 transitioning to State 1 or State 2 with probabilities ranging from 0.37 to 0.67. In contrast, the later stage exhibits a stabilized structure in which only States 1 and 2 persist, while the regression coefficients of States 2 and 3 converge, indicating behavioral alignment as wayfinding progresses. This pattern aligns with the observed relationship between initial latent state distribution and mode of hub arrival. State 1 is predominantly observed among metro-arriving passengers, highlighting a strong dependence on structured spatial information likely facilitated by the relatively organized nature of metro access points. In contrast, State 2 dominates among passengers arriving by road and air, reflecting a greater adaptability to less structured and more heterogeneous access environments. State 3, acting as a transitory exploratory phase, is slightly more prevalent among road and air travelers during the early phase of wayfinding. However, as behavior stabilizes and familiarity increases, State 3 rapidly vanishes, with individuals transitioning to either State 1 or State 2, depending on their navigational strategy and environmental interaction.

These results substantiate a dynamic adaptation process, wherein initial exploratory behavior gives way to more stable and differentiated wayfinding strategies based on arrival context, spatial familiarity, and individual navigation preferences.

Based on the above analysis, the three latent states can be interpreted as follows, each representing a distinct behavioral mode in the pedestrian wayfinding process.

State 1: A cautious, visually driven behavior mode. This state reflects a strategy characterized by a strong reliance on dense spatial and visual information. During the early stages, individuals tend to favor spaces that are visually open, structurally explicit, and offer salient path cues. The preference for navigation aids and reliance on prior travel experience serve to mitigate environmental uncertainty. Over time, as familiarity accumulates, the strategy transitions toward simplified execution, with reduced dependence on spatial cues. This state typifies a structured, guidance-oriented behavior, often observed in passengers arriving via metro who benefit from clearly defined spatial hierarchies.

State 2: A two-phase adaptive behavior mode linked to heterogeneous travel purposes. Primarily observed among passengers with non-essential or diverse travel purposes, this state unfolds in two distinct phases: In the early stage, personal goals such as ticketing or shopping lead to preference locking, resulting in low sensitivity to both visual cues and shortest-distance optimization. Pedestrians in this phase exhibit a high tolerance for local obstacles and weak reliance on navigational aids or prior knowledge. In the later stage, visual sensitivity significantly increases. Pedestrians display strong preferences for spatial features such as long isovist perimeters and edge-like obstacles, while actively avoiding open visual spaces and highly clustered areas. This shift represents a micro-navigation strategy, wherein individuals leverage fragmented visual cues and complex spatial edges to converge efficiently on target destinations while maintaining minimal cognitive load–a behavior particularly salient in the absence of navigation assistance.

State 3: A transient exploratory mode prior to goal fixation. This state functions as a short-lived transitional phase predominantly occurring in the early decision stage. It is marked by weak coefficient magnitudes and low sensitivity to both distance and visual attributes, indicating a lack of clear navigational intent. Pedestrians in this state are in the process of environmental scanning or delayed goal confirmation. As decisions solidify, individuals rapidly transition out of this state toward either State 1 or State 2, as reflected by the high early transition probabilities. State 3 thus captures a pre-decisional exploration phase preceding the adoption of a structured or adaptive wayfinding strategy.

By analyzing the transition relationships among the three aforementioned states, we can preliminarily develop a state regulation strategy for passenger organization in transport hubs, consciously guiding newly arrived passenger groups (state 3) toward the other two states. This approach aims to achieve a multi-objective balance among peak flow mitigation, transfer efficiency enhancement, and commercial friendliness. At key decision nodes, adjustable environmental cues and service processes can be deployed to modify passengers' perceptual weighting of spatial elements, thereby increasing transition probabilities. For instance, when the objective is rapid passenger dispersal, passengers in state 3 should be preferentially guided toward state 1. Before they reach the first decision node, the shortest path can be made salient–for example, by using advance signage or floor arrows to provide distant visual anchors for the next decision node–while reducing the visual prominence of suboptimal routes. This strategy helps passengers with limited time before departure proceed quickly to boarding. For other passengers without navigation assistance, fragmented cues along the route can steadily guide them toward commercial or service facilities, diverting a portion of travelers with heterogeneous trip purposes directly into state 2, where task-oriented routes dominate. This not only helps absorb main flow and alleviate pressure on bottleneck nodes but also maintains a low cognitive-cost movement toward goal areas without disrupting their intended activities.

6. Conclusion

This study proposes an innovative methodological framework that integrates a multi-state Hidden Markov Model with a dynamic Logit model (DHMM-Logit) to capture the temporal evolution of pedestrian route choice decisions and underlying individual heterogeneity in a VR-simulated integrated transport hub environment. By leveraging non-immersive VR to precisely reconstruct heterogeneous pedestrian travel processes and combining spatial syntax indicators with visual semantic segmentation data, the proposed approach effectively integrates both explicit spatial structural features and implicit visual-cognitive factors influencing wayfinding behavior. Furthermore, the incorporation of an improved Viterbi decoding algorithm and high-order spatial semantic encoding of prior decision nodes elevates the model's predictive accuracy to 97.54%. These innovations address the limitations of conventional observational methods and static discrete choice models in modeling cognitive states, explicitly revealing the latent cognitive mechanisms within pedestrian decision dynamics and offering interpretable inputs for predicting pedestrian flow scenarios in hub operations.

Empirical results from the Shanghai Hongqiao Integrated Transport Hub demonstrate that pedestrian route choices are significantly influenced by latent cognitive states, which evolve dynamically throughout the journey. Specifically, individuals in different cognitive states exhibit markedly distinct sensitivities to spatial connectivity, visual openness, and navigational cues, underscoring the necessity of dynamic modeling to account for behavioral heterogeneity. Moreover, the study quantitatively confirms that preferences for spatial features (e.g., visual connectivity) vary significantly depending on trip purpose (e.g., shopping, ticketing, waiting) and navigation tool usage, highlighting the practical importance of implementing differentiated spatial organization strategies and targeted information dissemination in hub management. From a practical perspective, the findings offer actionable insights for transport hub operators, enabling the optimization of spatial layout and navigation infrastructure according to user group characteristics. Results indicate that incorporating behavioral cognitive states and travel purposes into hub design can significantly improve passenger flow efficiency and satisfaction, thereby enhancing the operational resilience and service quality of large-scale transportation infrastructure.

Despite its theoretical and practical contributions, this study has several limitations. First, although the sample size is sufficient for model estimation, it may not fully capture the broader population diversity within real-world transport hubs, potentially limiting generalizability. Future research should expand the sample scope and scale to validate and refine model applicability across diverse user groups. Second, this study characterizes the visual information available at decision points using metrics of visual depth and breadth; however, because the eyes can rotate independently of the head/face, these proxies cannot precisely locate the point of regard nor recover fixation sequences, dwell time, and other aspects of true attentional allocation. By contrast, eye tracking can precisely record gaze positions and timing, offering a more direct view of how passengers allocate attention to signage and gates. Future work should incorporate eye-tracking devices and include multiple gaze-based measures as covariates in the DHMM-Logit model, thereby strengthening causal identification and improving the precision of design recommendations for hub wayfinding and facility layout. Finally, although the model explicitly incorporates spatial and visual variables, certain unobserved psychological factors, such as emotional states or situational stress may also influence pedestrian decisions. Subsequent research could integrate physiological indicators or subjective assessments to more comprehensively capture the complexity of cognitive processes in pedestrian behavior.

In conclusion, this study presents a significant methodological advancement in modeling pedestrian wayfinding behavior within complex transport hubs, providing a robust theoretical foundation and practical roadmap for enhancing passenger experience and operational efficiency through dynamic spatial management and informed guidance strategies.

CRediT authorship contribution statement

Zhicheng Dai: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization; **Dewei Li:** Supervision, Funding acquisition; **Yan Feng:** Supervision, Conceptualization; **Chenyi Yang:** Formal analysis, Data curation.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is supported by the Fundamental Research Funds for the Central Universities (Grant No. 2025JBZX077, 2022JBQY006, 2024YJS081) and the National Natural Science Foundation of China (72471023). We also thank the China Scholarship Council for its financial contribution

Appendix A. Participant characteristics questionnaire items

Instruction (before experiment): Please complete this questionnaire based on your actual personal situation and your actual travel process and needs within the hub.

1. Have you ever taken a high-speed rail train at Shanghai Hongqiao Station?
2. How long before departure do you usually arrive at the station to wait for your train?
3. During your stay in the transport hub, do you collect a paper ticket at a self-service or staffed ticket machine?
4. Do you usually make purchases (e.g., beverages, meals, gifts) inside the hubs when traveling by high-speed railway?
5. Please outline your regular sequence of activities once inside the station. (e.g., security → ticket collection → waiting area → boarding).
6. Prior VR experimental experience (yes/no): whether the participant had previously taken part in VR-based behavioral experiments?

Post-trial survey:

1. Cybersickness during/after the session (yes/no): whether any motion-sickness-related symptoms (e.g., nausea, dizziness, oculomotor discomfort) were experienced during or immediately after the current session.

Appendix B. Transportation hub node degree distribution and participant inter-regional transition frequency statistics

This appendix presents the degree distribution data of regional nodes within the hub scenario (Fig. B.1) and the actual inter-regional transition frequencies of participant groups (Fig. B.2), to facilitate readers' understanding of the transportation hub layout used in this study.

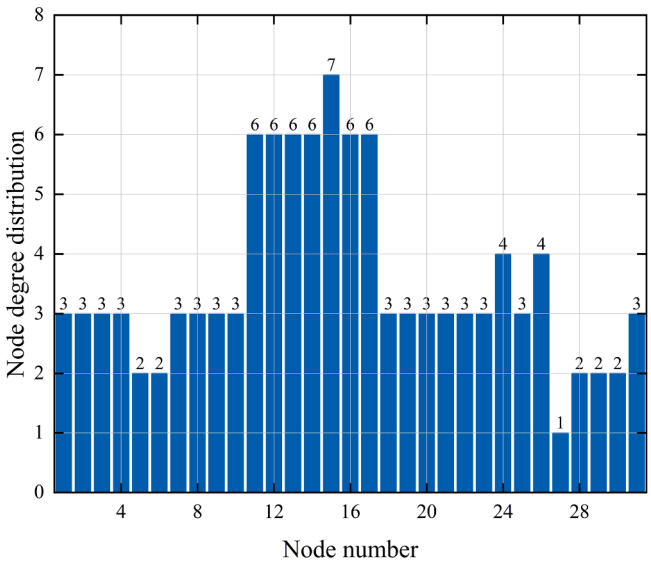


Fig. B.1. Statistical bar chart of degree distribution for nodes in transportation hub area.

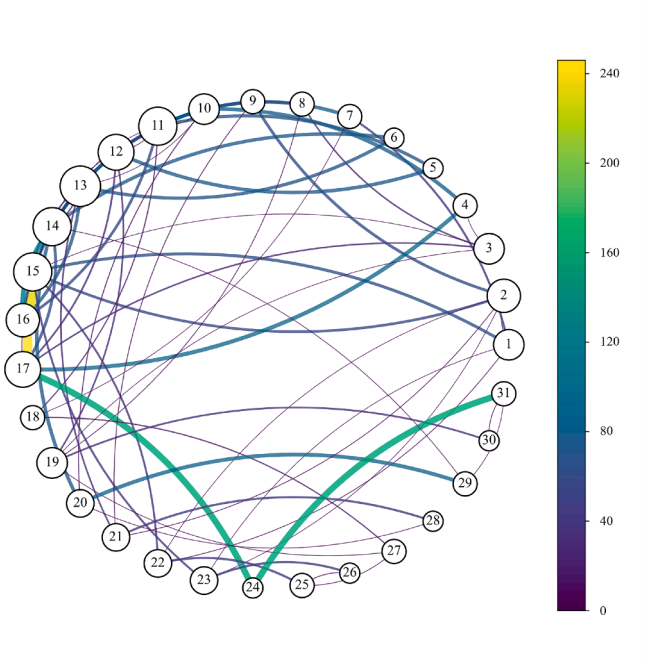


Fig. B.2. Passenger transition intensity network among hub areas. The line width and color represent different levels of transition frequency.

Appendix C. Sample size adequacy and statistical robustness testing

This appendix reports the test methods and operational procedures used for sample adequacy and statistical robustness testing. The results of the sample size adequacy and statistical robustness tests are presented in Table C.1.

1. Cluster Bootstrap by participant: Regression coefficients were re-estimated using participant-level resampling with replacement, and confidence intervals were constructed via the quantile method. This approach addresses the issue of variance underestimation caused by within-participant correlation across multiple sequences.
2. Permutation test: Within the participant set, treatment labels were randomly permuted while preserving the number of treated and control participants. The treatment effect was recalculated repeatedly to obtain a distribution free p-value, testing whether the observed effect could arise by chance under the current random assignment structure and finite sample conditions.
3. Monte Carlo Power: Given the observed group means, variances, and sample sizes, simulation was used to evaluate the probability of detecting a true effect at a specified significance level. This approach assesses whether the current sample size provides sufficient power to detect the main effects of heterogeneous factors.
4. Heterogeneity factor screening: For each participant, the number of steps to task completion and the cumulative local distance were computed as behavioral quantitative indicators. These measures were paired with individual heterogeneity attributes, and Welch's t-tests were conducted to examine the significance of behavioral differences across heterogeneous groups. The results are presented in Table C.2.

Table C.1
Results of sample size adequacy and statistical robustness tests.

Indicator	Bootstrap 95% CI	Permutation p-value	Power
Parameter value	$B_1 = 2000$	$B_2 = 1500$	$\alpha = 0.05$ Monte Carlo $B_3 = 5000$
Value	[97.26; 323.83]	0.0044	0.852

B_1- The number of resampling with replacement conducted using participants as sampling units.
B_2- Number of repetitions for the randomization permutation test
B_3- Number of sample groups generated by Monte Carlo efficacy analysis simulation.

Table C.2
Significance results of the influence of participant heterogeneity attributes on pathfinding behavior based on Welch's t-test.

Variable name	Statistical measure	p-value
2D Navigation	3.84	3.41×10^{-27}
Scenario familiarity	2.97	7.76×10^{-4}
Time of early arrival	3.01	1.48×10^{-15}
Gender	1.57	$0.184 > 0.05$

Appendix D. The detailed configuration of the Skip-gram model

The Skip-gram model used in this study employs Stochastic Gradient Descent (SGD) as the optimizer, with the negative sampling log-likelihood as the objective function. The initial learning rate was set to 0.025 and linearly decayed every 100 epochs to a final value of 0.0001. During tuning, model hyperparameters were selected based on the prediction accuracy and minimum BIC of the downstream DHMM-Logit model. Table D.3 provides detailed information on our model settings, and Fig. D.3 illustrates the impact of different hyperparameter configurations of the Skip-gram model on the performance of the DHMM-Logit model.

Table D.3
The detailed setting of the Skip-gram model configuration.

Module	Hyperparameters settings	Value
Skip-gram	Embedded Dimension	32
	Window size	5
	Epoch	300
	Learning rate	0.025 → 0.0001
	Optimizer	SGD
	Objective function	Log-likelihood of negative sampling

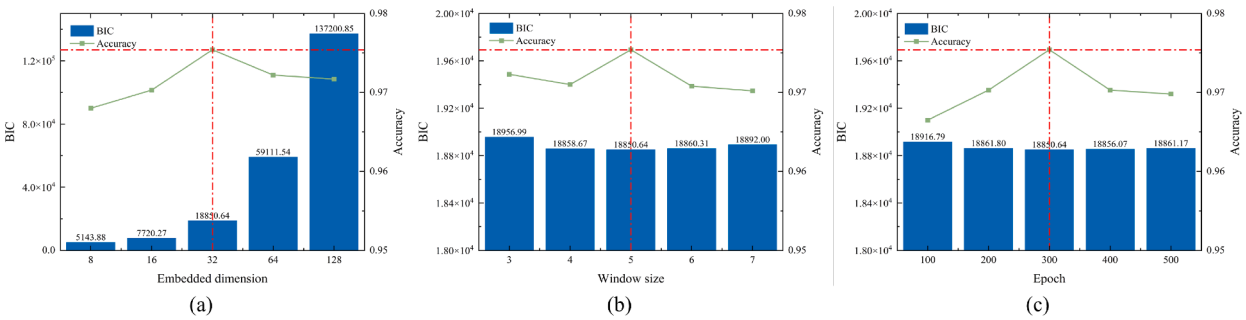


Fig. D.3. The performance of DHMM-Logit under different hyper-parameters of Skip-gram model.

References

Alinaghi, N., Giannopoulos, I., Kattenbeck, M., Raubal, M., 2025. Decoding wayfinding: analyzing wayfinding processes in the outdoor environment. Int. J. Geogr. Inf. Sci., 1–31, https://doi.org/10.1080/13658816.2025.2473599.

Andresen, E., Chraibi, M., Seyfried, A., 2018. A representation of partial spatial knowledge: a cognitive map approach for evacuation simulations. Transp. A Transp. Sci. 14 (5-6), 433–467. https://doi.org/10.1080/23249935.2018.1432717.

Asano, M., Iryo, T., Kuwahara, M., 2010. Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour. Transp. Res. Part C Emerging Technol. 18 (6), 842–855. https://doi.org/10.1016/j.trc.2010.01.005.

Barati, N., Zadegan, S. A.H., Kasravi, R., 2021. The role of survey details for wayfinding problem in complex pedestrian underground interchange with poor architectural configuration. Tunnelling Underground Space Technol. 108, 103718. https://doi.org/10.1016/j.tust.2020.103718.

Basu, R., Sevtsuk, A., 2022. How do street attributes affect willingness-to-walk? city-wide pedestrian route choice analysis using big data from boston and san francisco. Transp. Res. Part A Policy Pract. 163, 1–19. https://doi.org/10.1016/j.tra.2022.06.007.

Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., Ratti, C., 2011. Real-time urban monitoring using cell phones: a case study in rome. IEEE Trans. Intell. Transp. Syst. 12 (1), 141–151. https://doi.org/10.1109/TITS.2010.2074196.

Chen, N., Zhao, M., Gao, K., Zhao, J., 2020. The physiological experimental study on the effect of different color of safety signs on a virtual subway fire escape–an exploratory case study of zijing mountain subway station. Int. J. Environ. Res. Public Health 17 (16), 5903. https://doi.org/10.3390/ijerph17165903.

Cosma, G., Ronchi, E., Nilsson, D., 2016. Way-finding lighting systems for rail tunnel evacuation: a virtual reality experiment with oculus rift®. J. Transp. Saf. Secur. 8 (sup1), 101–117. https://doi.org/10.1080/19439962.2015.1046621.

Crociani, L., Vizzari, G., Yanagisawa, D., Nishinari, K., Bandini, S., 2016. Route choice in pedestrian simulation: design and evaluation of a model based on empirical observations. Intell. Artif. 10 (2), 163–182. https://doi.org/10.3233/IA-160102.

Dai, Z., Li, D., 2024. A research on pedestrian wayfinding behavior in large public space utilizing non-immersive VR platform. Collective Dyn. 9, 1–10. https://doi.org/10.17815/CD.2024.167.

Dai, Z., Li, D., Feng, Y., Yang, Y., Sun, L., 2024. A study of pedestrian wayfinding behavior based on desktop VR considering both spatial knowledge and visual information. Transp. Res. Part C Emerging Technol. 163, 104651. https://doi.org/10.1016/j.trc.2024.104651.

Dong, H., Zhou, M., Wang, Q., Yang, X., Wang, F.-Y., 2020. State-of-the-art pedestrian and evacuation dynamics. IEEE Trans. Intell. Transp. Syst. 21 (5), 1849–1866. https://doi.org/10.1109/TITS.2019.2915014.

Dong, W., Qin, T., Yang, T., Liao, H., Liu, B., Meng, L., Liu, Y., 2022. Wayfinding behavior and spatial knowledge acquisition: are they the same in virtual reality and in real-world environments? Ann. Am. Assoc. Geogr. 112 (1), 226–246. https://doi.org/10.1080/24694452.2021.1894088.

Dong, W., Wu, Y., Qin, T., Bian, X., Zhao, Y., He, Y., Xu, Y., Yu, C., 2021. What is the difference between augmented reality and 2D navigation electronic maps in pedestrian wayfinding? Cartogr. Geogr. Inf. Sci. 48 (3), 225–240. https://doi.org/10.1080/15230406.2021.1871646.

Dubey, R.K., Thrash, T., Kapadia, M., Hoelscher, C., Schinazi, V.R., 2019. Information theoretic model to simulate agent-signage interaction for wayfinding. Cognit. Comput. 13 (1), 189–206. https://doi.org/10.1007/s12559-019-09689-1.

Duives, D.C., Wang, G., Kim, J., 2019. Forecasting pedestrian movements using recurrent neural networks: an application of crowd monitoring data. Sensors 19 (2), 382. https://doi.org/10.3390/s19020382.

Dulebenets, M.A., Abioye, O.F., Ozguven, E.E., Moses, R., Boot, W.R., Sando, T., 2019. Development of statistical models for improving efficiency of emergency evacuation in areas with vulnerable population. Reliab. Eng. Syst. Saf. 182, 233–249. https://doi.org/10.1016/j.ress.2018.09.021.

Fang, Z., Song, W., Zhang, J., Wu, H., 2010. Experiment and modeling of exit-selecting behaviors during a building evacuation. Physica A 389 (4), 815–824. https://doi.org/10.1016/j.physa.2009.10.019.

Feng, Y., Duives, D., Daamen, W., Hoogendoorn, S., 2021. Data collection methods for studying pedestrian behaviour: a systematic review. Build. Environ. 187, 107329. https://doi.org/10.1016/j.buildenv.2020.107329.

Feng, Y., Duives, D.C., Hoogendoorn, S.P., 2022a. Development and evaluation of a VR research tool to study wayfinding behaviour in a multi-story building. Saf. Sci. 147, 105573. https://doi.org/10.1016/j.ssci.2021.105573.

Feng, Y., Duives, D.C., Hoogendoorn, S.P., 2022b. Wayfinding behaviour in a multi-level building: a comparative study of HMD VR and desktop VR. Adv. Eng. Inf. 51, 101475. https://doi.org/10.1016/j.aei.2021.101475.

Feng, Y., Xu, Z., Farah, H., Van Arem, B., 2024. Does another pedestrian matter? a virtual reality study on the interaction between multiple pedestrians and autonomous vehicles in shared space. IEEE Trans. Intell. Transp. Syst. https://doi.org/10.31219/osf.io/r3udx.

Haghani, M., Sarvi, M., 2017a. Following the crowd or avoiding it? empirical investigation of imitative behaviour in emergency escape of human crowds. Anim. Behav. 124, 47–56. https://doi.org/10.1016/j.anbehav.2016.11.024.

Haghani, M., Sarvi, M., 2017b. Stated and revealed exit choices of pedestrian crowd evacuees. Transp. Res. Part B Methodol. 95, 238–259. https://doi.org/10.1016/j.trb.2016.10.019.

Haghani, M., Sarvi, M., 2018. Crowd behaviour and motion: empirical methods. Transp. Res. Part B Methodol. 107, 253–294. https://doi.org/10.1016/j.trb.2017.06.017.

He, M., Wang, Q., Chen, J., Xu, S., Ma, J., 2023. Modeling pedestrian walking behavior in the flow field with moving walkways. Physica A 619, 128726. https://doi.org/10.1016/j.physa.2023.128726.

Hoogendoorn, S.P., 2004. Walking behavior in bottlenecks and its implications for capacity. In: 83rd Annual Meeting Transportation Research Board. Transportation Research Board (TRB), 1–13.

Jin, C.-J., Wu, C., Song, Y., Liu, T., Li, D., Jiang, R., Fang, S., 2024. The route choices of pedestrians under crowded and non-emergency conditions: two-route experiments and modeling. J. Choice Model. 50, 100463. https://doi.org/10.1016/j.jocm.2023.100463.

Kessler, F., Frankenstein, J., Rothkopf, C.A., 2024. Human navigation strategies and their errors result from dynamic interactions of spatial uncertainties. Nat. Commun. 15 (1), 5677. https://doi.org/10.1038/s41467-024-49722-y.

Kinateder, M., Comunale, B., Warren, W.H., 2018. Exit choice in an emergency evacuation scenario is influenced by exit familiarity and neighbor behavior. Saf. Sci. 106, 170–175. https://doi.org/10.1016/j.ssci.2018.03.015.

Kobes, M., Helsloot, I., De Vries, B., Post, J.G., Oberijé, N., Groenewegen, K., 2010. Way finding during fire evacuation; an analysis of unannounced fire drills in a hotel at night. Build. Environ. 45 (3), 537–548. https://doi.org/10.1016/j.buildenv.2009.07.004.

Kwee-Meier, S.T., Mertens, A., Jeschke, S., 2019. Recommendations for the design of digital escape route signage from an age-differentiated experimental study. Fire Saf. J. 110, 102888. https://doi.org/10.1016/j.firesaf.2019.102888.

Lcvenshtcin, V.I., 1966. Binary coors capable or 'correcting deletions, insertions, and reversals. In: Soviet Physics Doklady. 10, 8, 707–710.

Li, H., Zhang, J., Xia, L., Song, W., Bode, N. W.F., 2019. Comparing the route-choice behavior of pedestrians around obstacles in a virtual experiment and a field study. Transp. Res. Part C Emerging Technol. 107, 120–136. https://doi.org/10.1016/j.trc.2019.08.012.

Lieu, S.J., Guhathakurta, S., 2025. Exploring pedestrian route choice preferences by demographic groups: analysis of street attributes in chicago. Transp. Res. Part A Policy Pract. 195, 104437. https://doi.org/10.1016/j.tra.2025.104437.

Liu, R., Neisch, P., 2023. The impact of street infrastructure design on transit users' propensity to walk: a visual route choice experiment. Available at SSRN 4365991 https://doi.org/10.2139/ssrn.4365991.

Murtagh, E.M., Mair, J.L., Aguiar, E., Tudor-Locke, C., Murphy, M.H., 2021. Outdoor walking speeds of apparently healthy adults: a systematic review and meta-analysis. Sports Med. 51 (1), 125–141. https://doi.org/10.1007/s40279-020-01351-3.

Nilsson, D., Johansson, A., 2009. Social influence during the initial phase of a fire evacuation–analysis of evacuation experiments in a cinema theatre. Fire Saf. J. 44 (1), 71–79. https://doi.org/10.1016/j.firesaf.2008.03.008.

Perozzi, B., Al-Rfou, R., Skiena, S., 2014. DeepWalk: online learning of social representations. In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 701–710. https://doi.org/10.1145/2623330.2623732.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
Shen, H., 2019. Mutual information scaling and expressive power of sequence models. arXiv:1905.04271.
Srinivasan, A.R., Karan, F. S.N., Chakraborty, S., 2017. Pedestrian dynamics with explicit sharing of exit choice during egress through a long corridor. Physica A 468, 770–782. https://doi.org/10.1016/j.physa.2016.11.118.
Tong, Y., Bode, N. W.F., 2024. How building layout properties influence pedestrian route choice and route recall. Transp. A Transp. Sci. 20 (2), 2143249. https://doi.org/10.1080/23249935.2022.2143249.
van der Hoeven, F., van Nes, A., 2014. Improving the design of urban underground space in metro stations using the space syntax methodology. Tunnelling Underground Space Technol. 40, 64–74. https://doi.org/10.1016/j.tust.2013.09.007.
Wang, B., Ye, M., Li, X., Zhao, F., Ding, J., 2011. Abnormal crowd behavior detection using high-frequency and spatio-temporal features. Mach. Vis. Appl. 23 (3), 501–511. https://doi.org/10.1007/s00138-011-0341-0.
Wen, X., Si, B., Xu, M., Zhao, F., Jiang, R., 2024. A passenger flow spatial–temporal distribution model for a passenger transit hub considering node queuing. Transp. Res. Part C Emerging Technol. 163, 104640. https://doi.org/10.1016/j.trc.2024.104640.
Wong, L.T., Lo, K.C., 2007. Experimental study on visibility of exit signs in buildings. Build. Environ. 42 (4), 1836–1842. https://doi.org/10.1016/j.buildenv.2006.02.011.
Xu, Z., Wu, D., Yu, C., Chu, X., Sang, N., Gao, C., 2024. SCTNet: single-branch CNN with transformer semantic information for real-time segmentation. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38, pp. 6378–6386. https://doi.org/10.1609/aaai.v38i6.28457.
Yang, H.-D., Sin, B.-K., Lee, S.-W., 2003. Automatic pedestrian detection and tracking for real-time video surveillance. In: International Conference on Audio-and Video-Based Biometric Person Authentication. Springer, pp. 242–250. https://doi.org/10.1007/3-540-44887-X_29.
Yin, H., Cherchi, E., 2024. Preferences for automated taxis. a comparison between immersive virtual reality and screen-based stated choice experiments. Transp. Res. Part C Emerging Technol. 163, 104628. https://doi.org/10.1016/j.trc.2024.104628.
Yu, J., Mo, D., Zhu, Z., Chen, X.M., 2023a. A high-order hidden markov model for dynamic decision analysis of multi-homing ride-sourcing drivers. Transp. Res. Part C Emerging Technol. 148, 104031. https://doi.org/10.1016/j.trc.2023.104031.
Yu, L., Wang, Z., Chen, F., Li, Y., Wang, W., 2023b. Subway passengers' wayfinding behaviors when exposed to signage: an experimental study in virtual reality with eye-tracker. Saf. Sci. 162, 106096. https://doi.org/10.1016/j.ssci.2023.106096.
Yue, H., Zhang, B.-Y., Shao, C.-F., Xing, Y., 2014. Exit selection strategy in pedestrian evacuation simulation with multi-exits. Chin. Phys. B 23 (5), 050512. https://doi.org/10.1088/1674-1056/23/5/050512.
Zhou, Z.-X., Nakanishi, W., Asakura, Y., 2021. Route choice in the pedestrian evacuation: microscopic formulation based on visual information. Physica A 562, 125313. https://doi.org/10.1016/j.physa.2020.125313.